Indian Journal of Computer Science and Technology
https://www.doi.org/10.59256/indjcst.20260503130
Volume 5, Issue3 (September-December 2026), PP: 1087-1093.
www.indjcst.com
ISSN No: 2583-5300

ML-Based Multi-Customer Handwritten Signature Verification and Forgery Detection System

Swarnalatha[1], Sindhu K S[2], Chethan Kumar G S[3], Prasanna Kumar Naik R[4], Pallavi Pawar A[5]

[1,2,3,4] Assistant Professor, Department of Computer Applications, University BDT College of Engineering (A Constituent College of VTU, Belagavi), Davangere, Karnataka, India.

[5]Student, Department of Computer Applications, University BDT College of Engineering (A Constituent College of VTU, Belagavi), Davangere, Karnataka, India.

To Cite this Article: *Swarnalatha[1], Sindhu K S[2], Chethan Kumar G S[3], Prasanna Kumar Naik R[4], Pallavi Pawar A[5], "ML-Based Multi-Customer Handwritten Signature Verification and Forgery Detection System", Indian Journal of Computer Science and Technology, Volume 05, Issue 03 (September-December 2026), PP: 1087-1093.*

Copyright: ©2026 This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution License; Which Permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract: *Signatures continue to be used as identity evidence in financial transactions, legal records, insurance procedures and public services. Checking paper signatures manually can require considerable time, and the outcome may vary between examiners. The difficulty increases when a forged sample closely imitates the writer's usual appearance. This paper outlines an offline signature-verification framework designed to handle multiple customers through one machine-learning model. A writer-independent Siamese convolutional neural network (CNN) converts each signature image into a numerical representation, called an embedding. Representations from the same writer are intended to be close in the learned feature space, whereas representations from different writers or forged samples should be separated. During enrolment, a small collection of genuine samples is processed and its embeddings are stored for the relevant customer. A later query is assessed against the references belonging to the identity being claimed; adding another customer therefore requires reference enrolment rather than a new classifier. The proposed workflow covers image preparation, pair-based model learning and evaluation using false acceptance rate (FAR), false rejection rate (FRR), equal error rate (EER) and accuracy. A Histogram of Oriented Gradients (HOG) with Support Vector Machine (SVM) baseline is included for comparison. The paper also considers practical deployment, privacy safeguards and the need for human review.*

Key Word: *Offline Signature Verification, Handwriting Biometrics, Forgery Detection, Siamese CNN, Metric Learning, Customer Enrolment, Machine Learning.*

I. INTRODUCTION

Offline handwritten signature verification examines a static image and determines whether it is consistent with the identity stated by the person presenting it. This task is relevant to documents such as bank cheques, account forms, payment instructions and agreements. When large numbers of records must be checked, relying entirely on manual comparison can create delays and may lead to inconsistent decisions. A scalable automated method should therefore support a large customer base without requiring a separately trained classifier for every individual.

Signature systems generally operate in one of two settings. An online system records information while a person writes, including movement, timing and, where available, pen pressure. An offline system receives only a completed signature captured on paper by scanning or photography [11]. The latter setting is challenging because the sequence and dynamics of the writing process cannot be recovered directly from the image. Forgery attempts are commonly described as random, simple or skilled. In a random attempt, the writer has little or no knowledge of the target signature; a simple attempt may use only the target person's name; a skilled attempt is made after observing and practising the genuine signature. Skilled imitation is particularly difficult to detect because its overall shape may resemble the original closely.

The work aims to develop a reusable verification approach for a population of customers. The same trained feature extractor is intended to be shared across identities, while customer-specific reference embeddings are maintained separately. The output is a similarity-based decision that can assist a trained employee rather than replace all forms of document examination.

1.1 Motivation

The motivation for this work arises from several operational needs. Automated comparison may help flag suspicious documents before a transaction is completed. It can also reduce the time spent on repetitive first-level checking and apply a consistent scoring procedure to incoming samples. A shared model is attractive in settings where many customers must be supported but only a limited number of genuine signatures can be collected from each person. Finally, storing scores and decision thresholds can support later auditing, provided that biometric information is handled under suitable privacy and access controls.

1.2 Problem Statement

Let the service maintain N enrolled customers, with k genuine reference samples associated with each customer. For a new image q presented under the claimed identity c, the system must estimate whether q is consistent with that customer's references and return both a decision and a similarity score. The model should continue to support the existing customer population when a new identity is enrolled; retraining a separate classifier for that identity is not part of the intended workflow.

1.3 Contributions

- The framework uses one writer-independent feature-learning model to compare signatures across the enrolled customer population.
- Customer enrolment is based on storing embeddings generated from a small set of genuine reference images, rather than fitting a new classifier for each person.
- The evaluation design separates training, validation and test writers and reports errors for both random and skilled forgery cases using FAR, FRR and EER.
- The system design also considers customer-reference storage, selection of operating thresholds, performance monitoring and referral of uncertain cases to a human examiner.

1.4 Organisation of the Paper

The remainder of the paper is arranged as follows. Section II summarizes relevant research. Section III introduces the benchmark datasets and image preparation steps. Section IV explains the implementation sequence, followed by the architecture in Section V. The conventional approach and the proposed learning-based method are discussed in Sections VI and VII. Section VIII sets out the evaluation procedure. Deployment considerations, application areas, limitations and possible extensions are covered in Sections IX to XI, and Section XII closes the paper.

II. LITERATURE SURVEY

Prior work on signature verification includes feature-engineering methods as well as neural approaches that learn representations from examples. The following publications provide the technical background for the approach considered here.

1. Plamondon and Srihari [11] surveyed handwriting recognition and discussed both online and offline processing. Their work is useful for distinguishing systems that retain writing dynamics from those that operate on a final image, and for understanding the terminology used in this field.
2. Bromley and colleagues [1] described a Siamese time-delay neural network for comparing signature patterns. The central idea is to process two inputs through branches with shared parameters and use the resulting representations to estimate whether the inputs correspond. This pair-comparison formulation is an important foundation for writer-independent verification.
3. Hadsell, Chopra and LeCun [10] presented a contrastive objective for learning representations in which related examples are encouraged to occupy nearby positions and unrelated examples are separated by a margin. Such an objective is suitable for training networks that compare pairs of signature images.
4. Dalal and Triggs [8] introduced the HOG descriptor for representing local gradient direction patterns. Although it was developed for human detection, gradient histograms can also serve as hand-crafted image features in a conventional signature-classification pipeline, including an SVM baseline.
5. Hafemann, Sabourin and Oliveira [2] investigated deep convolutional features for offline handwritten signature verification. Their later review [3] discusses the progress and continuing difficulties in this area, including the limited number of genuine examples available for an individual writer and the challenge of distinguishing skilled imitations.
6. Dey et al. [4] proposed SigNet, which applies a convolutional Siamese architecture to writer-independent offline signature comparison. Their experiments include established collections such as CEDAR [5] and BHSig260 [7], making the work relevant to the dataset choices discussed in this paper.
7. Koch, Zemel and Salakhutdinov [9] explored Siamese networks for one-shot image recognition. Their study provides motivation for learning a comparison function that can be applied to a previously unseen identity using only a limited set of reference examples.

Approach	Features	Classifier / Matching	Main limitation
Writer-dependent SVM	HOG, LBP, geometric	One SVM per customer	Needs many samples; does not scale
Distance statistics [5]	Gradient, structural, concavity	Distance to references	Hand-crafted; weaker on skilled forgeries
Deep features + SVM [2]	Learned by CNN	Writer-dependent SVM on CNN features	Still one classifier per user
Siamese CNN [1], [4]	Learned embedding	Distance between embedding's	Needs pair mining and threshold tuning

Table 1: Qualitative comparison of signature verification approaches

Taken together, these studies support the use of learned representations when the task requires comparison across writers. Hand-crafted descriptors remain useful as interpretable and computationally straightforward baselines, while a writer-independent metric-learning model offers a natural way to add identities without fitting a new classifier for every customer. Performance still depends on the diversity and quality of training data, pair construction and threshold calibration.

III. DATASETS AND PER-PROCESSING

3.1 Benchmark Datasets

The datasets used in offline signature research vary in language, writer population, image characteristics and the availability of skilled imitations. Using more than one collection can help reveal whether a method generalizes beyond the conditions represented in a single benchmark. The dataset figures summarized in Table 2 should be checked against the original dataset publications before final submission.

Dataset	Writers	Signatures per writer	Notes
CEDAR [5]	55	24 genuine, 24 skilled forgeries	English signatures; small, widely used
BHSig260 [7]	260	24 genuine, 30 forgeries	100 Bengali and 160 Hindi writers
GPDS Synthetic [6]	Thousands (synthetic)	Genuine and skilled forgeries	Large scale; synthetically generated

Table 2: Public datasets for offline signature verification (confirm exact figures against each dataset's documentation)

3.2 Image Pre-processing

Signature scans may differ in dimensions, background condition, ink appearance and stroke width. The preparation stage standardizes these inputs so that differences caused by scanning conditions have less influence on the learned representation. The proposed sequence is outlined below.

- Convert each input to a single-channel gray scale image so that the model receives a consistent intensity representation.
- Apply a median or Gaussian smoothing operation, where appropriate, to reduce isolated speckles and acquisition noise.
- Use Otsu thresholding to separate foreground strokes from the background, then invert the result when needed so that the signature strokes are represented with higher intensity.
- Locate the foreground region and trim unnecessary blank margins around the handwritten mark.
- Preserve the original width-to-height relationship by padding the cropped image before resizing it to the selected network input dimensions, for example 155 × 220 pixels.
- Scale pixel intensities to the interval from 0 to 1 to provide a consistent numerical input to the CNN.

3.3 Data Augmentation

Genuine examples from one writer are often limited, so the training set can be expanded with restrained image transformations. Possible operations include rotations of approximately ±5 degrees, modest shear, scale variation of about ±10%, small translations and controlled changes in stroke thickness through morphological dilation or erosion. These transformations should reproduce plausible capture or writing variation without altering the defining characteristics of the signature. Augmentation must be applied only to training data, not to validation or test samples.

IV. IMPLEMENTATION

The implementation can be organized into the stages described below, from preparing the source images to monitoring the deployed service.

- **Data collection and preparation:** The benchmark collections supply genuine and forged examples from multiple writers. In an operational setting, a customer would provide a small number of genuine specimens during enrolment; the proposed design assumes approximately three to five references. Each image is then processed using the steps described in Section III.
- **Pair construction:** Training examples are formed as pairs. A positive pair contains two genuine signatures from the same writer. Negative pairs may combine a genuine sample with a skilled imitation or compare genuine samples belonging to different writers. The pair-generation procedure should maintain a suitable balance between the two labels.
- **Feature extraction:** The CNN encoder applies convolutional operations together with batch normalization, ReLU activation and pooling. Its final fully connected layers map the input image to a 128-dimensional feature vector. This vector is the embedding used during subsequent comparisons.
- **Network learning:** Two copies of the encoder process the two images in a pair, but their weights are shared so that both inputs are represented in the same feature space. The model is optimized with the contrastive loss [10] and the Adam optimizer. Writers must be assigned to mutually exclusive training, validation and test partitions; otherwise, results may be inflated by the presence of familiar identities in the test set.
- **Customer enrolment and checking:** During enrolment, the system runs each approved reference image through the encoder and stores the resulting vectors under the customer identifier. For a later request, it computes the query vector and compares it only with the references linked to the identity being claimed.

- **Validation:** Testing on writers excluded from model development provides an estimate of writer-independent performance. Report accuracy, FAR, FRR, EER and ROC-AUC, and distinguish results for random attempts from those for skilled imitations.
- **Deployment and maintenance:** The trained encoder can be exposed through separate enrolment and verification services. The system should retain appropriate decision logs and monitor score behaviour over time. If operating conditions or signature patterns change, the threshold may need recalibration and the model may require further training using a properly governed dataset.

V. SYSTEM ARCHITECTURE

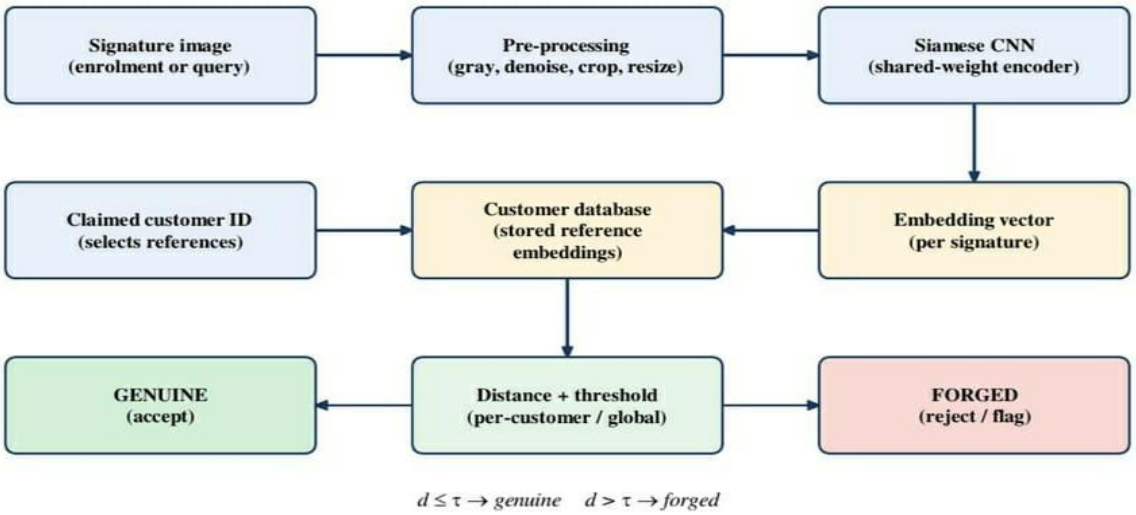


Figure 1: Processing architecture for the proposed multi-customer signature verification framework.

As illustrated in Figure 1, a reference image is prepared and passed through the shared CNN encoder when a customer is enrolled. The resulting feature vector is stored in the customer database. A submitted signature follows the same image-preparation and encoding path. The system retrieves the reference vectors associated with the claimed customer, calculates the distance between the query and those references, and compares the resulting score with a decision threshold τ. The threshold may be global or adjusted for an individual customer using the variation observed among that customer's enrolled samples.

VI. EXISTING MODEL

A conventional offline verification pipeline often begins by calculating manually designed descriptors from each signature. Examples include HOG [8], local binary patterns, geometric measurements and features calculated over image grids. These descriptors can then be supplied to an SVM. In a writer-dependent arrangement, a separate classifier is trained for each customer using genuine samples and suitable negative examples. This approach is relatively direct and can be effective when enough labelled examples are available for every identity.

There are, however, practical weaknesses in this arrangement. Designed descriptors may fail to represent subtle stroke differences that are important when the imitation is carefully prepared. Maintaining an independent classifier for every customer also increases the training and administration burden as the customer base grows. In addition, only a small number of genuine samples may be available for a new customer, making a customer-specific classifier vulnerable to overfitting. Genuine enrolment generally does not provide a collection of that customer's forged signatures, which further complicates supervised training.

VII. PROPOSED MODEL

The proposed approach uses a writer-independent Siamese CNN [1], [4] to learn how to compare two signature images. Instead of assigning every image to a fixed customer class, the network produces a feature representation that can be compared with stored examples. Since the encoder is shared across writers, the same trained parameters can be used for different identities. Enrolling another customer involves generating and storing that person's reference embeddings, rather than training a new customer-specific model.

7.1 Encoder and Loss

Let $f(\cdot)$ represent the CNN encoder, and let x_1 and x_2 denote two pre-processed signature images. Their embedding distance is $D = ||f(x_1) - f(x_2)||_2$. The pair label is $y = 0$ when the pair is genuine and $y = 1$ when it represents a non-matching or forged pair. The contrastive objective [10] used in this design is given below:

$$L = (1 - y) \cdot D^2 + y \cdot max(0, m - D)^2$$

Here, m denotes the separation margin. For a genuine pair, the objective penalizes a large distance between the two embedding's. For a negative pair, it applies a penalty when the distance remains below the selected margin. In this way,

training encourages the network to form compact representations for matching examples and more separated representations for non-matching examples.

7.2 Encoder Architecture

The shared encoder is based on a compact convolutional design influenced by earlier Siamese signature-verification work [4]. Table 3 lists an initial configuration for experimentation. The layer sizes and regularization settings are proposed starting values and should be validated empirically rather than treated as universally optimal.

Layer	Configuration	Output
Input	1 × 155 × 220 gray scale image	Image
Conv1 + BN + ReLU + MaxPool	96 filters, 11 × 11; pool 3 × 3, stride 2	Feature maps
Conv2 + BN + ReLU + MaxPool	256 filters, 5 × 5; pool 3 × 3, stride 2; dropout 0.3	Feature maps
Conv3 + BN + ReLU	384 filters, 3 × 3	Feature maps
Conv4 + BN + ReLU + MaxPool	256 filters, 3 × 3; pool 3 × 3, stride 2; dropout 0.3	Feature maps
FC1 + ReLU	1024 units; dropout 0.5	Vector
FC2 (embedding)	128 units, L2-normalised	Embedding

Table 3: Proposed Siamese encoder configuration (shared by both branches)

7.3 Training Configuration

Parameter	Proposed value
Optimiser	Adam, learning rate 1e-4 with decay on plateau
Loss	Contrastive loss, margin m = 1.0
Batch size / epochs	64 pairs / up to 50 epochs with early stopping on validation EER
Embedding size	128
Reference signatures per customer (k)	3 to 5 genuine samples
Data split	Customer-disjoint train / validation / test

Table 4: Proposed training configuration

7.4 Verification Algorithm

For a query image q submitted under customer identity c, the verification routine proceeds as follows:

- Step 1: Load the k reference embeddings stored for customer c.
- Step 2: Apply the defined image-preparation procedure to q and obtain its embedding, f(q).
- Step 3: Measure the distance between f(q) and every reference embedding for the selected customer.
- Step 4: Combine the individual distances, using a preselected rule such as their mean or minimum, to obtain one comparison score d.
- Step 5: Compare d with threshold τ. A score at or below the threshold is classified as genuine; a larger score is classified as a possible forgery.
- Step 6: Return the classification and comparison score, and record the result in the system log.

7.5 Threshold Selection

The value of τ determines the balance between false acceptance and false rejection. It may be selected using validation data, for example at the equal-error operating point. For transactions where accepting a forgery would have a particularly high cost, the operating point may instead be chosen to reduce FAR, with uncertain cases sent for manual examination.

VIII. EXPERIMENTAL SETUP AND EVALUATION PLAN

Evaluation is planned on public signature datasets using a writer-disjoint division of the data. The principal measures used to describe verification performance are defined below.

- FAR is calculated as the number of forged test samples incorrectly accepted divided by the total number of forged samples tested.
- FRR is the number of genuine test samples incorrectly rejected divided by the total number of genuine samples tested.
- EER is the operating error value at the point where the false-acceptance and false-rejection rates are equal.
- Accuracy is the proportion of all test decisions that are correct; ROC-AUC may additionally be reported to describe discrimination across thresholds.

Model	Accuracy (%)	FAR (%)	FRR (%)	EER (%)
HOG + SVM (baseline)	[to be measured]	[to be measured]	[to be measured]	[to be measured]
Siamese CNN (proposed)	[to be measured]	[to be measured]	[to be measured]	[to be measured]

Table 5: Performance comparison template (insert the values obtained from your own experiments)

For a meaningful comparison, results should identify the dataset, the number of enrolled references per writer and the distance-aggregation rule. Random and skilled forgeries should be reported separately because they represent different levels of difficulty. Further experiments can vary the number of references k, the embedding dimension and the contrastive-loss margin to examine their effect on performance.

IX. APPLICATIONS AND SYSTEM DEPLOYMENT

9.1 Application Areas

- Banking and financial services: preliminary screening of signatures on cheques, withdrawal instructions, transfer documents, account applications and loan paperwork.
- Insurance and legal administration: comparison of signatures appearing on policy documents, claim forms, contracts, wills and authorization records.
- Public and educational institutions: checking signatures submitted with certificates, applications, official forms and attendance documents.
- Forensic document examination: supplying a repeatable numerical comparison that may assist, but should not replace, an expert's assessment.

9.2 Multi-Customer Database Design

The customer record can contain an identifier, the embeddings produced from approved reference signatures, the enrolment date and, where justified, a customer-specific threshold. Keeping feature vectors for routine comparison can reduce the amount of data that must be accessed during verification because only the query image needs to pass through the encoder. If original signature images are retained for audit or legal purposes, they should be encrypted and made available only to authorized personnel.

9.3 Operational Workflow

- Enrolment: obtain k verified genuine specimens, prepare each image, calculate its embedding and associate the vectors with the customer record.
- Verification: accept a query image and claimed customer identifier, calculate its embedding-based score, apply the relevant threshold and return the outcome.
- Human review: direct cases with scores near the decision boundary to an authorized examiner for a second assessment.
- Reference maintenance: when a customer's genuine writing pattern changes, add new samples only after their authenticity has been independently confirmed.

9.4 Security, Privacy and Ethical Considerations

Signature images and their derived representations are biometric information and require careful handling. Collection should be based on an appropriate legal basis and clear notice or consent where required. Data should be protected during transfer and storage, access should be restricted, and retention periods should be defined. An automated score should not be the only reason for an adverse decision: a review mechanism and a way to challenge the outcome are important. Evaluation should also examine performance across writing scripts, capture conditions and relevant user groups to identify uneven error patterns.

X. ADVANTAGES, CHALLENGES AND LIMITATIONS

10.1 Advantages

- A shared feature extractor can support a growing customer population without fitting an independent model for every new enrollee.
- The enrolment process can operate with a limited number of genuine samples and does not require the customer to supply forged examples.
- The network learns image representations from training examples, potentially capturing stroke patterns that fixed hand-designed descriptors may not express adequately.
- Once reference embedding's have been stored, verification requires encoding the query and comparing it with the selected customer's reference set.

10.2 Challenges and Limitations

- Carefully practised imitations remain difficult to distinguish because their visible shape may closely follow the genuine signature.
- A single person's signatures are not identical on every occasion; natural changes may be associated with writing conditions, age, health, writing instrument or surface.

- Poor scans, low resolution, stamps, printed lines and overlapping marks can obscure strokes or introduce misleading image features.
- A model developed using one script or acquisition environment may not transfer reliably to another. New conditions may require threshold adjustment and further validation.
- Benchmark datasets may not represent the full range of real-world fraud. Their size, collection procedure and controlled forgery conditions can limit how confidently results transfer to operational use.

XI. FUTURE SCOPE

- Integrate static signature images with dynamic measurements such as pen pressure, writing speed and pen angle when suitable digital capture devices are available.
- Investigate attention-based networks, vision transformers and alternative metric-learning objectives to improve the quality of learned signature representations.
- Explore few-shot and meta-learning approaches for situations in which enrolment provides only one or two genuine examples.
- Test performance across different writing systems and languages, particularly where a service must process signatures from multilingual populations.
- Extend the threat model to include digitally manipulated or machine-generated imitations and examine resistance to adversarial input changes.
- Develop interpretable outputs that indicate which image regions or stroke patterns contributed to a comparison score, while avoiding unsupported claims about explainability.
- Study model compression and quantization so that verification can be performed on mobile or edge hardware with limited computing resources.

XII. CONCLUSION

This paper has described a proposed framework for verifying offline handwritten signatures belonging to multiple customers. Its central component is a writer-independent Siamese CNN that maps signature images into an embedding space for similarity comparison. Customer-specific reference vectors are stored at enrolment, allowing a newly registered identity to use the existing encoder without fitting a separate classifier. The paper has set out the relevant datasets, image preparation operations, pair-generation strategy, network configuration, verification procedure and evaluation measures. A HOG–SVM method is included as a conventional point of comparison. The framework is intended to support document-screening workflows, not to replace expert judgement in difficult cases. Before practical use, the proposed configuration must be trained and tested on appropriately separated writers, with measured results reported for both random and skilled forgeries. Data protection, threshold selection and access to human review are also necessary parts of a responsible deployment.

REFERENCES

1. J. Bromley, I. Guyon, Y. LeCun, E. Säckinger and R. Shah, “Signature verification using a Siamese time delay neural network,” Advances in Neural Information Processing Systems (NIPS), 1993.
2. L. G. Hafemann, R. Sabourin and L. S. Oliveira, “Learning features for offline handwritten signature verification using deep convolutional neural networks,” Pattern Recognition, vol. 70, 2017.
3. L. G. Hafemann, R. Sabourin and L. S. Oliveira, “Offline handwritten signature verification – Literature review,” Int. Conf. on Image Processing Theory, Tools and Applications (IPTA), 2017.
4. S. Dey, A. Dutta, J. I. Toledo, S. K. Ghosh, J. Llados and U. Pal, “SigNet: Convolutional Siamese network for writer independent offline signature verification,” arXiv:1707.02131, 2017.
5. M. K. Kalera, S. Srihari and A. Xu, “Offline signature verification and identification using distance statistics,” Int. J. Pattern Recognition and Artificial Intelligence, vol. 18, no. 7, 2004. (CEDAR dataset)
6. M. A. Ferrer, M. Diaz-Cabrera and A. Morales, “Static signature synthesis: A neuromotor inspired approach for biometrics,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 37, no. 3, 2015. (GPDS Synthetic)
7. S. Pal, A. Alaei, U. Pal and M. Blumenstein, “Performance of an off-line signature verification method based on texture features on a large Indic-script signature dataset,” IAPR Workshop on Document Analysis Systems (DAS), 2016. (BHSig260 dataset)
8. N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2005.
9. G. Koch, R. Zemel and R. Salakhutdinov, “Siamese neural networks for one-shot image recognition,” ICML Deep Learning Workshop, 2015.
10. R. Hadsell, S. Chopra and Y. LeCun, “Dimensionality reduction by learning an invariant mapping,” IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2006.
11. R. Plamondon and S. N. Srihari, “Online and off-line handwriting recognition: A comprehensive survey,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 1, 2000.