International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 11 | Nov - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

Virtual Police Station: An AI-Assisted Platform for Online FIR, Emergency SOS and Police Station Management

Sonawane Vaishnavi Shamsundar[1], Madane Shweta Ganesh[2], Dolse Vaishnavi Ganesh[3], Garje Ankita Babasaheb[4], Prof. P.V.Gaikwad[5], Prof. H. B. Jadhav[6]

[1]Student, Department of Computer Engineering, Adsul's Technical Campus, Chas, Ahilyanagar, India, [sonawanevaishnavi220@gmail.com]
[2]Student, Department of Computer Engineering, Adsul's Technical Campus, Chas, Ahilyanagar, India, [madaneshweta04@gmail.com]
[3]Student, Department of Computer Engineering, Adsul's Technical Campus, Chas, Ahilyanagar, India, [dolasevaishanvi@gmail.com]
[4]Student, Department of Computer Engineering, Adsul's Technical Campus, Chas, Ahilyanagar, India, [ankitagarje@gmail.com]
[5]Guide, Department of Computer Engineering, Adsul's Technical Campus, Chas, Ahilyanagar, India, [pratiksha.atc@gmail.com]
[6]HOD, Department of Computer Engineering, Adsul's Technical Campus, Chas, Ahilyanagar, India, [hemantjadhav090@gmail.com]

Abstract - Most police services in India still require the citizen to visit the station, both to report a crime and to learn what happened to the complaint, while the station keeps complaints, FIRs, criminal records, lock-up entries, evidence and staff details in separate registers. Section 173 of the Bharatiya Nagarik Suraksha Sanhita, 2023 now allows information about a cognizable offence to be given electronically and makes Zero FIR statutory. This paper presents a Virtual Police Station that serves citizens, police staff and administrators from one platform. Citizens register with OTP and Aadhaar e-KYC, file complaints and e-FIRs in English, Marathi or Hindi by text or voice, track cases through five stages and raise a one-tap SOS that sends live location to the nearest station, with an SMS fallback for weak networks. An AI service ranks complaints by urgency with a fine-tuned MuRIL classifier and keyword rules, flags duplicate complaints with a score that combines Sentence-BERT similarity with the distance and time between incidents, matches suspect photos against criminal profiles using MTCNN and ArcFace embeddings, and transcribes spoken complaints with Whisper. Police staff register FIRs with an HMAC-signed QR code and manage cases, lock-up, evidence custody and staff duty, and all sensitive access is written to a hash-chained audit log. The system is built with React, React Native, Node.js, MongoDB and FastAPI, and is evaluated on a team-built trilingual complaint dataset, the LFW face benchmark and SOS delivery tests on 4G and 2G networks.

Key Words: virtual police station, e-FIR, Zero FIR, emergency SOS, duplicate detection, urgency classification, face recognition.

1. INTRODUCTION

A police station does two kinds of work for the public. It receives information about crimes and emergencies, and it records what it does about them: FIRs, investigations, arrests, persons in lock-up, seized property and the officers on duty. In most districts both kinds of work still happen at a desk inside the station. A citizen whose phone was stolen on a highway may have to travel to the station that has jurisdiction, a woman may not be comfortable describing harassment at a public counter, and a senior citizen may not be able to travel at all. After the complaint is given, there is usually no simple way to know whether an FIR was registered or who is investigating, except by visiting again.

Inside the station the problems are of a different kind. Complaints are read in the order they arrive, so a complaint about a threat with a weapon can wait behind several reports of lost documents. Two witnesses of the same accident may file two complaints that become two FIRs. Matching a suspect photograph against old records depends on an officer remembering faces, and lock-up, evidence and duty registers kept on paper are hard to audit.

The legal framework now supports a digital process. Section 173 of the Bharatiya Nagarik Suraksha Sanhita (BNSS), 2023, in force from 1 July 2024, allows information to be given by electronic communication, to be taken on record once the informant signs it within three days, and requires the station to record information irrespective of the area where the offence was committed [12]. State citizen portals already offer a few services, and the Crime and Criminal Tracking Network and Systems (CCTNS) links station records [14], but the citizen side, the station side and AI support are rarely designed as one system.

This work designs and builds such a system. The main contributions are: (i) a three-role platform covering complaint, e-FIR and Zero FIR filing, case tracking, SOS, criminal records, lock-up, evidence custody, staff management and administration; (ii) an urgency score that combines a MuRIL classifier with keyword, category and vulnerability signals for complaints written in English, Marathi and Hindi; (iii) a duplicate score that adds the distance and time between incidents to Sentence-BERT text similarity, with a geospatial pre-filter that keeps the check fast; (iv) an SOS path that stays under 1 KB per message and falls back to SMS; and (v) a security design with two-factor login, role-based access control, field encryption, QR-verified FIR documents and a hash-chained audit log.

The rest of the paper is organised as follows. Sec. 2 reviews related work. Sec. 3 describes the proposed system. Sec. 4 gives the experimental setup, Sec. 5 presents the results and discussion, and Sec. 6 concludes the paper.

2. RELATED WORK

2.1 Online Crime Reporting

Several student and state projects provide an online complaint form with a status page, usually built as a web application over a relational database. They move the first step online but leave verification, case management and station registers outside the system, and they do not help the officer decide which complaint to read first. The BNSS provisions on e-FIR and Zero FIR [12] and the Digital Personal Data Protection Act, 2023 [13] set new requirements on signing, jurisdiction-free recording and consent for personal data that such systems must now follow.

2.2 Face Recognition

FaceNet [1] showed that a network can map faces to an embedding space where distance measures identity, so new people can be added without retraining. ArcFace [2] added an angular margin to the training loss and gave more separable 512-dimensional embeddings, reaching 99.83% verification accuracy on LFW [11]. MTCNN [3] detects faces and five landmarks with a cascade of three small networks and is commonly used to align faces before recognition.

2.3 Text Similarity, Classification and Speech

BERT [5] introduced pre-training of a bidirectional Transformer followed by task fine-tuning. Sentence-BERT [4] fine-tuned BERT in a siamese structure so that sentences are compared by cosine similarity, reducing a search over 10,000 sentences from hours to seconds. MuRIL [6] pre-trained BERT on 17 Indian languages including transliterated text, which matters because many users type Marathi and Hindi in Roman letters. Whisper [7] transcribes speech in many languages, including Hindi and Marathi, without task-specific training.

2.4 Access Control, Logging and Crime Mapping

Role-based access control [8] assigns permissions to roles rather than users and supports role hierarchies, which matches the rank structure of a police station. Crosby and Wallach [10] showed how hash-linked logs make changes to past entries detectable. Chainey et al. [9] compared hotspot mapping methods on real crime data and found kernel density estimation the most useful for predicting where crimes occur.

3. PROPOSED SYSTEM

3.1 System Architecture

Fig. 1 shows the architecture. Citizens use a React Native mobile app or a React web portal, and police and admin users use a React portal. All traffic passes through an Nginx gateway that terminates TLS 1.3 [16] and limits the request rate. The application layer is written in Node.js with Express and is split into services for authentication, complaints and FIRs, cases and station records, SOS, and administration. MongoDB stores all records as JSON documents with 2dsphere indexes on station, incident and SOS locations; Redis holds OTPs and BullMQ job queues for SMS, PDF and embedding work; MinIO stores evidence files encrypted with AES-256. A FastAPI service hosts the four AI models and is reachable only from the application layer.

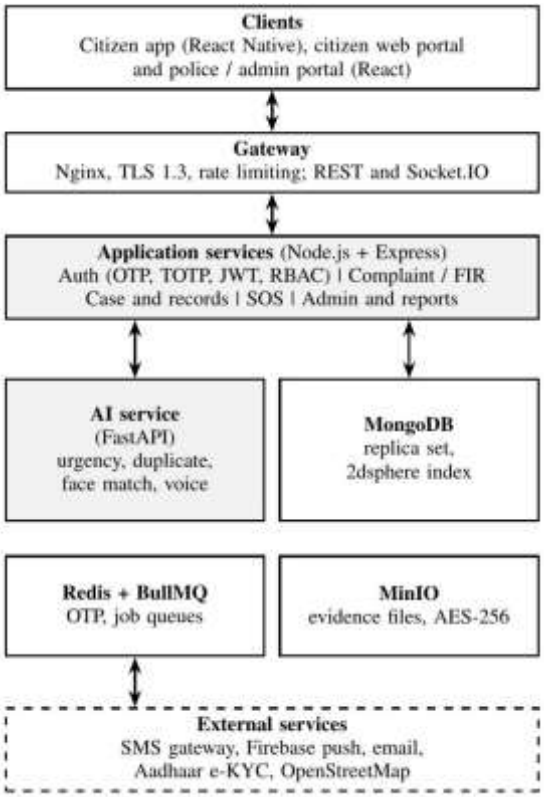


Fig -1: System architecture of the Virtual Police Station

3.2 Modules and User Roles

Table 1 lists the modules available to each role. Permissions are checked on the server for every request against a role and rank table, so a constable can record lock-up entries and evidence handovers, an inspector can register FIRs and file chargesheets, and the station house officer approves leave and duty swaps.

Table -1: Main modules of each role

Role	Modules
Citizen	OTP and 2FA login, Aadhaar e-KYC; complaint, e-FIR, Zero FIR and anonymous tip with evidence and voice entry; case tracking; SOS and emergency contacts; station rating; area alerts; missing and wanted lists
Police staff	Complaint queue with urgency and duplicate flags; FIR registration and transfer; case assignment, updates and chargesheet; criminal records with face search and vehicle search; lock-up; evidence with QR tags and custody chain; attendance, leave, duty swap, fatigue alerts
Admin	Dashboard of all stations; users and roles; stations and categories; crime heatmap and trends; ratings; audit logs; monthly reports

3.3 Urgency Ranking and Duplicate Detection

When a complaint is saved, the complaint service finds the station whose boundary polygon contains the incident location and sends the complaint to the AI service, which runs two checks in parallel (Fig. 2). For urgency, MuRIL [6] is fine-tuned with four output classes, and the probability of the urgent class p_u is combined with three rule-based signals:

$$P = 0.5\, p_u + 0.2\, s_{kw} + 0.2\, s_c + 0.1\, s_v \qquad (1)$$

where s_{kw} is a keyword score from a list of critical words in the three languages (weapon, kidnap, bleeding, threat and their equivalents), s_c is a weight for the category and s_v marks a vulnerable person such as a child or a senior citizen. Complaints with $P \geq 0.8$ are Critical, 0.6 to 0.8 High, 0.35 to 0.6 Medium and the rest Low. The station queue is ordered by P and then by filing time, and the officer can change the level. The keyword list is a configuration value, so a new term can be added without retraining.

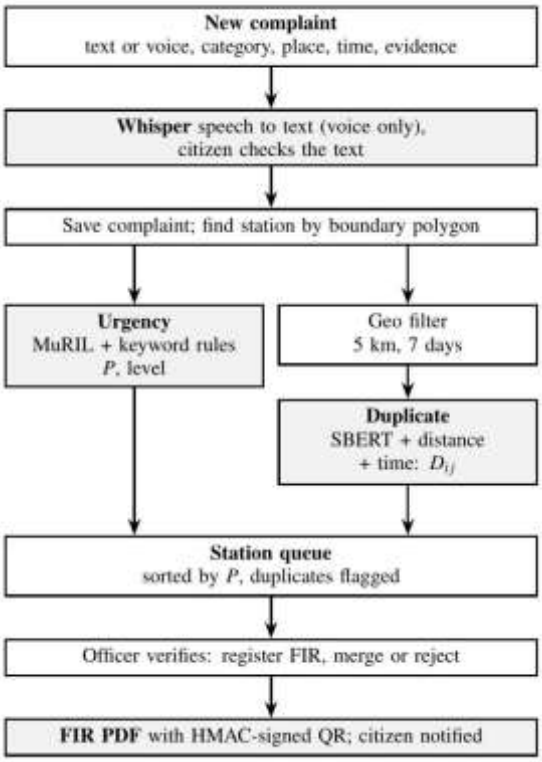


Fig -2: Processing of one complaint

Text similarity alone cannot separate two chain snatching cases reported on the same road a week apart. For duplicates, a 2dsphere query first selects stored complaints within 5 km and 7 days of the new one. Each candidate j is then scored against the new complaint i:

$$D_{ij} = 0.5 \cos(e_i, e_j) + 0.2\, e^{-d/\sigma_d} + 0.2\, e^{-|\Delta\tau|/\sigma_t} + 0.1\, a_{ij} \quad (2)$$

where e_i and e_j are 384-dimensional multilingual Sentence-BERT [4] embeddings, d is the distance between the incidents, $\Delta\tau$ the time difference, $\sigma_d = 1$ km, $\sigma_t = 24$ h and a_{ij} is 1 when the accused name or phone number matches. A warning is shown when $D_{ij} \geq 0.75$, and the officer decides whether to merge or keep the complaint.

3.4 Face Matching and Voice Entry

An officer can upload a photo to search the criminal database. MTCNN [3] detects and aligns the face, and ArcFace [2] maps it to a 512-dimensional embedding q. Each stored profile has an embedding g_i computed when the profile photo is saved, and the similarity is:

$$s_i = (q \cdot g_i) / (\|q\| \, \|g_i\|) \quad (3)$$

The five profiles with the highest s_i above a threshold θ_f are shown with their scores. The system never links a person to a case on its own; the officer must confirm the match. For voice entry, the app records up to three minutes of audio, the Whisper [7] small model transcribes it, and the citizen corrects the text before submission.

3.5 Emergency SOS

Fig. 3 shows the SOS path. A tap on the SOS button, or a shake where the phone allows it, sends a message with the user ID, latitude, longitude, accuracy, time and battery level, which is under 1 KB. With mobile data the message goes over Socket.IO and repeats every 10 seconds; without data the app sends an SMS with a map link to the gateway number and to the emergency contacts. The SOS service finds the nearest station with a $nearSphere query; the distance it uses is the haversine formula [17]:

$$d = 2R \arcsin \sqrt{[\sin^2(\Delta\varphi/2) + \cos\varphi_1 \cos\varphi_2 \sin^2(\Delta\lambda/2)]} \quad (4)$$

with R = 6371 km. All on-duty officers of that station see the live path on a map, and the first officer to respond is recorded.

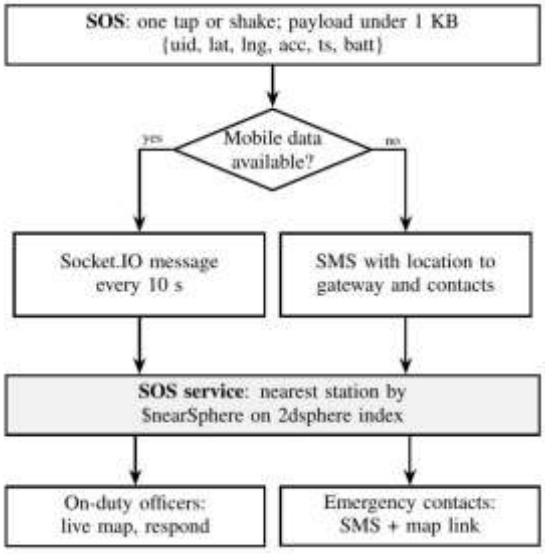


Fig -3: SOS delivery with SMS fallback

3.6 Security and Data Protection

Citizens log in with an OTP that is valid for five minutes; police and admin users also enter a TOTP code [15], and all users are logged out after 15 minutes of inactivity. Passwords are stored as bcrypt hashes and sessions use short-lived JWT access tokens. Aadhaar numbers, face embeddings and evidence files are encrypted with AES-256-GCM, Aadhaar is shown only in masked form, and consent is recorded before any Aadhaar, face or fingerprint data is collected, as required by [13]. Each FIR PDF carries a QR code with h = HMAC-SHA256(k_s, FIR number, station, date, content hash), which the public verification page recomputes. Every view or change of a sensitive record is appended to an audit log in which each entry stores the hash of the previous one:

$$H_i = \text{SHA256}(H_{i-1} \,\|\, e_i) \quad (5)$$

so that editing or deleting a past entry breaks the chain [10]. The database runs as a three-node replica set with an automatic encrypted backup every day at 02:00, kept for 30 days.

4. EXPERIMENTAL SETUP

4.1 Datasets

Real complaint texts are confidential, so a trilingual dataset is prepared by the team. About 3,000 complaints in English, Marathi and Hindi (in both Devanagari and Roman script) are written from public news reports with names and places changed, covering theft, assault, cyber crime, missing person, lost item and harassment. Each complaint is labelled with one of four urgency levels by two team members, with disagreements settled by the guide. For duplicates, 400 pairs are built in which the same incident is described by different people, together with hard negative pairs of similar incidents at different places or times. Face matching is evaluated on the LFW benchmark [11], and voice entry on 150 recorded complaints read by volunteers in the three languages. The dataset is split into 70% training, 15% validation and 15% test sets.

4.2 Evaluation Metrics

Urgency classification and duplicate detection are measured with precision, recall and F1 on the test set:

$$F1 = 2 \cdot \text{Precision} \cdot \text{Recall} / (\text{Precision} + \text{Recall}) \quad (6)$$

Urgency is reported as macro F1 over the four levels and separately as recall of the Critical class, because missing a

critical complaint costs more than a false alarm. Face matching is reported as the true accept rate at a false accept rate of 0.1% on LFW, and as top-5 accuracy on a gallery of 1,000 identities. Voice entry is measured by word error rate, WER = (S + D + I) / N, where S, D and I are substituted, deleted and inserted words and N the number of reference words. SOS delivery time is measured from the button press to the alert on the station screen, 50 times each on 4G, on a 2G connection and with data turned off.

4.3 Implementation Details

The backend uses Node.js 20, Express 4, Socket.IO, Mongoose, MongoDB 7, Redis 7 and MinIO. The clients use React 18 with Material UI, React Native with Expo, Leaflet with OpenStreetMap tiles and i18next for English, Marathi and Hindi. The AI service uses Python 3.11, FastAPI, PyTorch, Hugging Face Transformers, sentence-transformers, facenet-pytorch for MTCNN, InsightFace for ArcFace and Whisper. MuRIL is fine-tuned on a Google Colab T4 GPU for 4 epochs with AdamW, a learning rate of 2×10^{-5} and a batch size of 16; inference runs on CPU. SMS is sent through MSG91 and push notifications through Firebase Cloud Messaging.

5. RESULTS AND DISCUSSION

Table 2 gives the scores of the AI components on the test set and the measured SOS delivery time. Higher precision, recall, F1 and accuracy and lower WER and delivery time are better.

Table -2: Evaluation of the AI components and SOS delivery

Task	Method	P	R	F1
Urgency	Keyword rules	[ ]	[ ]	[ ]
	TF-IDF + logistic regression	[ ]	[ ]	[ ]
	MuRIL + rules (proposed)	[ ]	[ ]	[ ]
Duplicate	Text similarity only	[ ]	[ ]	[ ]
	Text + distance + time (proposed)	[ ]	[ ]	[ ]
Face match	ArcFace, TAR at FAR 0.1%; top-5 accuracy	[ ]	[ ]	[ ]
Voice entry	Whisper small, WER (En / Mr / Hi)	[ ]	[ ]	[ ]
SOS delivery	Seconds on 4G / 2G / SMS	[ ]	[ ]	[ ]

[Fill Table 2 with the values obtained from the test runs and describe them here: the gain of MuRIL with rules over keyword rules and TF-IDF, the recall of the Critical class, how many false duplicate warnings are removed by adding distance and time, the face match accuracy on LFW and on the gallery, the WER for each language, and the median and 95th percentile SOS delivery time on each network.]

Some behaviour can be checked on the running system independent of the averages. Complaints written in Roman-script Marathi or Hindi are the hardest case for keyword rules, which is where the MuRIL model is expected to help most. The geo and time terms in the duplicate score are expected to remove most false warnings between similar crimes at different places, but two genuinely separate incidents at the same place on the same day will still be flagged, which is why the final decision stays with the officer. Face matching depends heavily on photo quality: old, low-resolution or side-view photographs reduce the similarity of true matches, so the system shows five candidates rather than one. The SMS fallback makes the SOS slower than the data path but keeps it working when only a voice network is available. A limitation of the current work is that the complaint dataset is written by the team, so the urgency classifier should be retrained on real anonymised complaints before any field use.

6. CONCLUSION

This paper presented a Virtual Police Station that lets citizens file complaints, e-FIRs and Zero FIRs online in English, Marathi or Hindi, track their cases and raise SOS alerts that work on weak networks, while police staff and administrators manage FIRs, cases, criminal records, lock-up, evidence and staff in one system. AI support ranks complaints by urgency, flags duplicates using text, distance and time together, suggests matching criminal profiles from a photo and converts speech to text, while every legal decision stays with the officer. Two-factor login, role-based access control, encryption, QR-verified FIRs and a hash-chained audit log protect the data. Future work includes data exchange with CCTNS, integration with the 112 emergency response system, digital signatures on e-FIRs, support for more Indian languages and a guided chatbot for complaint filing.

ACKNOWLEDGEMENT

We thank our guide, Prof. P.V.Gaikwad., for guidance throughout this work, and the Department of Computer Engineering, Adsul's Technical Campus, Chas, Ahilyanagar, for providing facilities. We also thank the volunteers who recorded voice complaints for testing.

REFERENCES

1. F. Schroff, D. Kalenichenko and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering," in Proc. IEEE CVPR, 2015, pp. 815–823.
2. J. Deng, J. Guo, N. Xue and S. Zafeiriou, "ArcFace: Additive Angular Margin Loss for Deep Face Recognition," in Proc. IEEE/CVF CVPR, 2019, pp. 4690–4699.
3. K. Zhang, Z. Zhang, Z. Li and Y. Qiao, "Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks," IEEE Signal Processing Letters, vol. 23, no. 10, pp. 1499–1503, 2016.
4. N. Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," in Proc. EMNLP-IJCNLP, 2019, pp. 3982–3992.
5. J. Devlin, M.-W. Chang, K. Lee and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in Proc. NAACL-HLT, 2019, pp. 4171–4186.
6. S. Khanuja et al., "MuRIL: Multilingual Representations for Indian Languages," arXiv:2103.10730, 2021.
7. A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey and I. Sutskever, "Robust Speech Recognition via Large-Scale Weak Supervision," in Proc. ICML, PMLR vol. 202, 2023, pp. 28492–28518.
8. R. S. Sandhu, E. J. Coyne, H. L. Feinstein and C. E. Youman, "Role-Based Access Control Models," IEEE Computer, vol. 29, no. 2, pp. 38–47, 1996.
9. S. Chainey, L. Tompson and S. Uhlig, "The Utility of Hotspot Mapping for Predicting Spatial Patterns of Crime," Security Journal, vol. 21, pp. 4–28, 2008.
10. S. A. Crosby and D. S. Wallach, "Efficient Data Structures for Tamper-Evident Logging," in Proc. USENIX Security Symposium, 2009, pp. 317–334.
11. G. B. Huang, M. Ramesh, T. Berg and E. Learned-Miller, "Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments," Univ. of Massachusetts, Amherst, Tech. Rep. 07-49, 2007.
12. Government of India, The Bharatiya Nagarik Suraksha Sanhita, 2023 (Act No. 46 of 2023), Section 173.
13. Government of India, The Digital Personal Data Protection Act, 2023 (Act No. 22 of 2023).
14. National Crime Records Bureau, Ministry of Home Affairs, "Crime and Criminal Tracking Network and Systems (CCTNS)." [Online]. Available: https:ncrb.gov.in
15. D. M'Raihi, S. Machani, M. Pei and J. Rydell, "TOTP: Time-Based One-Time Password Algorithm," RFC 6238, IETF, 2011.
16. E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.3," RFC 8446, IETF, 2018.
17. R. W. Sinnott, "Virtues of the Haversine," Sky and Telescope, vol. 68, no. 2, p. 159, 1984.
18. N. Provos and D. Mazières, "A Future-Adaptable Password Scheme," in Proc. USENIX Annual Technical Conference, FREENIX Track, 1999.