Babak Abad – Artificial Intelligence and Robotics Specialist

09921687098 | Babak.Abad@yahoo.com | https://www.linkedin.com/in/babak-abad
www.ai-programmer.com | https://github.com/babak-abad | Iran – Tehran

10 years of experience in software development and the design of machine vision systems, natural language processing, image and text generation models. Capability in model optimization, debugging, and decomposing complex tasks into solvable and manageable components, resulting in rapid prototyping and the development of modules that are easily maintained.

Proficiency:	**Deep Learning**	**Machine Learning**	**Image Processing**	**Generative AI**
Natural Language Processing		**Data Fusion**	**Sensor Fusion**	**Augmentation**
Retrieval-Augmented Generation		**Edge Device**	**Transfer Learning**	**Optimization** .
Languages:	**Python**	**C++**	**C#**	**C (AVR)**
Tools:	**VS Code**	**PyCharm**	**Visual Studio**	**Colab**
Matlab Coder	**Matlab ToolBox**	**Docker**	**Git**	**Lens**
Netron	**AnyLabeling**	**CVAT**		

Work Experiences

AI Model Developer at Payam Gostar Tose'e Iranian (Tehran – 2024 / MAY to present).

- **Text Summarization**: Text embeddings were extracted using the BGE-M3 model and subsequently passed via LangChain to OpenAI's model, a process that proved not to be cost-effective. To optimize this procedure, after embedding extraction by BGE-M3, the embeddings were clustered using the OPTICS algorithm. Then, from the texts nearest to the centroid of each cluster, a small subset of samples was selected—rather than using the entire set of texts—and these samples were fed into the mT5 model for summary extraction.

- **Named Entity Recognition (NER):** A model leveraging XLM-RoBERTa communicates with a database via gRPC or Kafka and performs named entity recognition (NER) by employing two queues and the producer–consumer pattern. Through optimization of the number of phrases per batch, quantization of the model's layers using the Transformers library and PyTorch, and adjustment of the queue capacities, execution speed was improved by a factor of approximately two on CPU and three on GPU.

- **Development of a text detector based on YOLO 11 and CRNN:** Through transfer learning, the final layers of VGG Net were adapted to train a text detector within the Ultralytics framework. Cross-batch normalization layers mitigated VRAM constraints during training, enabling higher accuracy under limited computational resources. Hyperparameter optimization used the Optimum library. A CRNN combining convolutional and LSTM architectures was implemented in PyTorch to transcribe YOLO-detected text regions. Text was then organized into lines and paragraphs via OPTICS clustering. Data augmentation employed imgaug and torchvision. The vanishing feature map issue was addressed using pytorch-grad-cam and TorchCam (based on Grad-CAM). Rather than training on full Persian characters, glyphs were used as recognition units. Given the visual similarity among many Persian glyphs, each glyph was manually inspected and labeled, reducing learning time and required layers compared to the prior OCR architecture. The proposed method achieves 89% accuracy on a dataset of 20,000 Persian text images from social media.

- **Image search:** In this method, the CLIP model was employed for embedding extraction, and the resulting embeddings were stored in the Milvus database. Principal component analysis (PCA) was applied to reduce storage requirements and enhance retrieval speed. Moreover, PCA was periodically retrained—conditioned on elapsed time or the number of insertions—to mitigate the risk of learning locally effective vectors and to ensure that the resulting vectors remained representative of the entire database.

- **Speech-to-text:** The service retrieves a video or audio file from an S3 bucket, isolates the audio stream using the Librosa library, and subsequently transcribes the audio content by means of the Whisper model.

Software Engineer at Tajeran-e Jadeh Abrisham (Tehran – 2019 / Jun to 2024 / May)

- **Vehicle speed measurement via image processing:** In this approach, vehicle speed was initially estimated via license plate tracking. Detected plate texts and spatial coordinates were queued and clustered based on plate text similarity (Levenshtein and Hamming distances) and spatiotemporal distance, since plates are often illegible at a distance. Speed measurement error across sixteen cameras under varying conditions was 2%, as validated by traffic police records. An alternative method using vehicle tracking was then adopted. The DeepSORT ReID component was customized: trained on 100,000 images, replacing pairwise and Euclidean losses with a triplet approach using cosine distance. This reduced required cameras by 50% and enabled recording of vehicles without license plates.

- **License plate character recognition (OCR)**: A CRNN model was trained with Keras using Albumentations and AugLy for data augmentation. Under challenging conditions (shadows, motion blur, dust, spray, snow, distortion, low/high illumination), it achieved 99.4% accuracy (confirmed by traffic police). To address the recurrent layers' computational bottleneck, the architecture was redesigned as a LeNet-based model with batch processing and Bayesian optimization via Optimum and cppflow. This improved inference speed 22× with no accuracy loss and maintained compatibility with ONNX and protocol buffer (.pb) files from older software versions.

- **License plate localization:** Camera setup and frame capture used the Pylon library. YOLOv5-based plate detection with Bayesian optimization, DeepViz layer reduction, and OpenVINO integration boosted speed from 14 FPS to 60 FPS with >98% accuracy. Switching to YOLO removed the need for polarized and infrared-pass filters. Quantized layers enabled execution on a Raspberry Pi 4 + Neural Compute Stick 2, dropping accuracy to 92% but cutting power from 70W to 7W. With multi-stage verification, accuracy returned to 98%, suitable for low-speed applications like parking lot entrances.

- **Camera, radar, and flash synchronizer:** Initially, the FUTR3D model was employed to fuse data from heterogeneous sensors, including multiple cameras, radar, and LiDAR. Subsequently, to reduce system costs and address the unavailability of radar and LiDAR hardware, a bird's-eye-view (BEV) approach was adopted for data fusion. In addition, to synchronize the sensors with floodlights, a circuit based on an AVR microcontroller was designed. This implementation reduced projector operating temperatures by 80 degrees Celsius, extended floodlight operational lifespan from one year to over six years, and decreased electrical power consumption by a factor of twelve.

- **License plate generator:** A conditional generative adversarial network (C-GAN) was employed to generate license plate images under diverse visual conditions, thereby facilitating the evaluation or training of plate detection methods. The model generates plates exhibiting various attributes, including asymmetric illumination, Persian characters and digits, and multiple shadow types, among other variable parameters. Additionally, data augmentation techniques—such as noise injection, blurring, geometric stretching, piecwise deformation/transforms, rotation, and contrast adjustment—were applied to increase the diversity and volume of the generated samples.

- **Formulation of hardware requirements:** A set of parametric formulas was developed to determine the required specifications for computing hardware, cameras, lenses, floodlight configurations (number and type), and network interface cards as functions of maximum vehicle speed, traffic volume, road width, and application type—namely, speed measurement, red-light enforcement, toll collection, and no-entry monitoring. These formulas additionally incorporate camera parameters such as quantum efficiency, dynamic range, maximum signal-to-noise ratio, dark noise, rise time, frame rate, and pixel size, thereby ensuring qualitative compatibility when selecting new cameras relative to predefined minimum requirements.

Software Engineer at Dadeh Kavan-e Khazar Pouya (Rasht – 2016 / Dec to 2019 / Jun).

- **License plate detection:** Camera configuration and image acquisition were performed using the FLIR library. In this ad-hoc method, the vehicle license plate was located through edge detection and morphological operations, after which hyperparameters were optimized via Gaussian optimization within the Optimization Toolbox of MATLAB, yielding an accuracy of 92% (confirmed by the traffic police) and a speed of 5 FPS on a Core i7 2670QM.

- **License plate character reading (OCR):** A two-stage method combining thresholding and contour detection was employed. For thresholding, the adaptive Bradley method from the AForge.NET library was initially used; subsequently, to improve execution speed, the dlib implementation of the Bradley method was adopted. Contour

detection was carried out using the Emgu CV library (a wrapper for OpenCV). At this stage, geometric moments, Zernike moments, Hu moments, and Tchebichef moments were utilized as feature extraction methods. The classifiers—support vector machine (SVM) and perceptron—were selected based on analysis of results obtained from the AnyPredict and AutoKeras libraries.

- **Design of automatic labeling and reporting software:** The software provides a platform for annotating vehicle location, license plate text, characters and digits, plate type, vehicle class, vehicle class (sedan, heavy, trailer, etc), model (Peugeot 206, 405, Citroën Xantia, Renault L90, Toyota Camry, and 40 more), and vehicle color. To expedite the annotation workflow, previously trained models are leveraged for automatic labeling, and the resulting annotations are stored in a Microsoft SQL database. The reporting module evaluates the performance of each method based on license plate detection precision, plate reading accuracy, and the rates of missed plate types, employing metrics such as recall, receiver operating characteristic (ROC), precision, and false positive rate.

Software Developer at Avid (Rasht – 2013 / Apr to 2024 / Apr)

- **Customer relationship management system via tracking and identification:** Customer relationship management (CRM) in a chain store was implemented through indoor customer tracking to enable service functions, including guiding a customer who remained stationary at a given location, identifying specific and potential customers, and indicating available parking spaces. The software utilized ultra-high frequency radio-frequency identification (UHF-RFID) cards.

Education

Master's degree – Artificial Intelligence and Robotics Engineering (Azad University, Lahijan Branch – 2017 / Sep to 2019 / Sep)

- Thesis: Increasing the detection rate of Iranian vehicle license plates conforming to the European standard using Canny edge detection

Bachelor's degree – Software Engineering (University of Guilan –2008 / Sep to 2014 / Jan)

- Project: License plate detection using digital image processing techniques and artificial neural networks.

Publications

- B. A. Fomani and A. Shahbahrami, "***License plate detection using adaptive morphological closing and local adaptive thresholding***," 2017 3rd International Conference on Pattern Recognition and Image Analysis (IPRIA), Shahrekord, 2017, pp. 146-150, doi: 10.1109/PRIA.2017.7983035.

- Shahbahrami, B. A. Foomani and A. Akoushideh, "***A style-free and high speed algorithm for*** License ***Plate Detection,***" 2017 10th Iranian Conference on Machine Vision and Image Processing (MVIP), Isfahan, Iran, 2017, pp. 76-81, doi: 10.1109/IranianMVIP.2017.8342372.

- T. Najafi, B. A. Foomani and A.Shahbahrami, "***Anxiety and Depression Detection using Statistical Features,***" 2017 2nd National Conference on Soft Computing, Roudsar, Guilan, 2017.

- Pourkashani, A. Shahbahrami and B. A. Fomani, "***Copy Move Forgery Detection Using Histogram Quantization of Cross Power Spectrum,***" 2017 International Conference on High Performance Computing & Simulation (HPCS), Genoa, 2017, pp. 826-831, doi: 10.1109/HPCS.2017.125.

- B. A. Foomani, A. Shahbahrami and M. Ahmadi, "***Iranian Private Vehicle License Plate Detection using Image Processing Technics,***" 2013 8th Iranian Conference on Machine Vision and Image Processing (ICMVIP08), Iran, Zanjan, 2013.