

MODELO DE PROPOSTA DE PREÇOS

AO SETOR DE COTAÇÃO DO FESP

Segue abaixo, proposta comercial:

Item	Descrição	Unidade	QTD	Valor Total
01.	Aquisição de solução tecnológica especialista em reconhecimento de padrões biométricos faciais, corpóreos, identificação de veículos e placas. Software - Licença Perpétua.	Licença perpétua	02	
VALOR TOTAL:				

Empresa:

CNPJ:

Endereço:

Dados Bancários:

Telefone:

E-mail:

Prazo de entrega:

Validade da Proposta:

_____, _____ de _____ de 2025.

Assinatura do Representante da Empresa

O SOFTWARE DEVERÁ ATENDER OS REQUISITOS MÍNIMOS DESCRITO ABAIXO:

GERAL:

1. Software de sistema de reconhecimento facial, corpos, carros e placas veiculares baseado em inteligência artificial (sobre rede neural) com analíticos que funcione baseado em CPU e GPU, permitindo ainda trabalhar com múltiplas placas aceleradoras no mesmo servidor ou em múltiplos servidores com arquitetura escalável.
2. Detecção sem máscara de proteção respiratória com 50 pixels de largura no rosto para stream de vídeo.
3. Deve fazer as detecções e reconhecimento com máscara de proteção respiratória, com pelo menos 80 pixels de largura no rosto para stream de vídeo
4. Realizar carga de fotos (formatos webp, jpg, png, bmp) no sistema a partir de 60 pixels entre pupilas.
5. Necessário conseguir fazer o reconhecimento de qualquer tipo de tom de pele. Com a pessoa de lado 30 graus, com a pessoa com parte do rosto coberto por máscara de proteção respiratória também.
6. Deve detectar silhuetas para realizar contagem e distanciamento de entre pessoas, podendo criar a contagem a partir de uma única câmera ou múltiplas câmera dentro do mesmo contador.
7. Deve também realizar contagem de faces criando a mesma a partir de uma única câmera ou múltiplas câmera dentro do mesmo contador.
8. Ser capaz de definir uma região de interesse para o contador de faces e/ou silhuetas e se necessário desenhar poligonalmente a área de interesse.
9. Na mesma cena, o sistema deve ser capaz de detectar/ reconhecer no mínimo 40 faces com as mínimas condições de tamanho por face.
10. O fabricante deve ter histórico de implementação de um projeto com ao menos 100.000 licenças de reconhecimento facial.
11. Deve conseguir reconhecer corpos e fazer filtragens baseada em cor da roupa na parte superior e/ou inferior
12. Deve ser capaz de detectar carros, realizar filtragens por tipo de carroceria, fabricante, modelo, cor e placa
13. A interface gráfica do sistema deve suportar o idioma português
14. Ser capaz de aumentar e diminuir o “full frame” do evento selecionado com zoom através do mouse
14. Dispor de recurso que faça círculo de contato com pessoas marcadas como potenciais infectadas ou suspeitas de COVID-19, em dois níveis. A mesma versão deve permitir identificar aglomeração, distanciamento social e grupo de risco
16. Dispor de recurso que faça círculo de contato com pessoas marcadas como potenciais infratores, em três níveis
17. Necessário possuir o recurso de vivacidade (certificar-se de que é uma pessoa viva – “liveness”) para uso com stream proveniente de câmeras de CFTV
18. Dispor de recurso para integração com sistemas de controle de acesso. O conversor tcp/ip wiegand recebe do sistema de reconhecimento facial um webhook contendo o “facility code” e número do cartão da pessoa identificada, enviando essa informação através do protocolo Wiegand ao sistema de controle de acesso, para validação do acesso (ou não) da mesma.
19. Dispor da capacidade de integração com sistema de controle de acesso, em nível de software via API.
20. Possuir o recurso para postar apenas a melhor detecção, implica selecionar dentro de um conjunto de frames que formam a detecção (período de acompanhamento de uma pessoa em frente a câmera que está ativo, ou seja, enquanto a face estiver sendo

detectada pelo sistema em frente a câmera sem interrupção), o melhor frame em questão de qualidade para reconhecimento e descartar os demais.

21. Possuir a capacidade de permitir de-duplicação de detecções e reconhecimentos de uma mesma pessoa que passe em mais de uma câmera (configurada dentro de um mesmo grupo de câmeras) para gravação de eventos únicos dentro de intervalo de tempo pré-definido, mantendo apenas o evento de melhor qualidade.

22. Possuir o recurso para postar múltiplas capturas durante uma detecção, implica postar todas as detecções possíveis dos frames que formam a detecção (período de acompanhamento de uma pessoa em frente a câmera que está ativo, ou seja, enquanto a face estiver sendo detectada pelo sistema em frente a câmera sem interrupção).

23. Possuir recurso de verificação capaz de comparar faces, corpos e carros.

24. Possuir recurso de agrupamento de eventos por similaridade de vetores:

24.1. O sistema deve conseguir catalogar de forma única, cada indivíduo que se apresente em frente às câmeras do sistema;

24.2. A medida que o mesmo indivíduo apareça no vídeo das diferentes câmeras e em diferentes momentos, todos eventos devem ser agrupados dentro do mesmo catálogo do indivíduo criado inicialmente.

25. Dispor de recurso para monitorar em câmera(s), a presença de pessoas em uma área determinada, podendo serem criadas regras de:

25.1. Agendamento

25.2. Gatilhos para número mínimo de pessoas (a ser parametrizado), durante um período de tempo (a ser parametrizado) em dada área.

25.3. Gatilhos para número máximo de pessoas (a ser parametrizado), durante um período de tempo (a ser parametrizado) em dada área. 26. Dispor de mosaico de vídeo (videowall) capaz de monitorar até 9 câmeras simultâneas, sobrepondo aos objetos detectados (faces/corpos/carros) em tempo real:

26.1. Marcação da localização do objeto no frame (contorno de uma “caixa” retangular)

26.2. Atributos do objeto (Ex. facial – Idade, gênero, etc..)

ARQUITETURA:

27. Necessário trabalhar com arquitetura centralizada, distribuída ou híbrida.

27.1. Arquitetura local: Todo processamento é realizado localmente, desde a decodificação dos streams, detecção das faces e vetorizações faciais. A base de dados fica em cada servidor sendo gerenciada de forma independente.

27.2. Arquitetura centralizada: sendo todos os streams chegando a um servidor central e processados neste ambiente, onde estarão base de dados, aplicação e decodificação.

27.3. Arquitetura distribuída: podendo ter parte da aplicação na borda (decodificação de vídeo), enviando apenas as detecções com as imagens já normalizadas para um servidor de aplicação central (aplicação) e servidor de armazenamento central (base de dados).

28. A interface gráfica do sistema deve ser baseada em web (“web client”), funcionando nos principais navegadores de mercado, sem requerer a instalação de nenhum programa adicional

29. O software do sistema deverá suportar o sistema operacional Linux Ubuntu

30. O sistema deverá ter a capacidade de operar com uma base de dados de no mínimo 100 milhões de pessoas de interesse

31. Precisa suportar streaming de vídeos nos padrões HTTP e RTSP compressão H.264, MPEG em 25 fps e resolução 1080p com bit rate mínimo de 4Mb/s.

32. Deve possuir suporte ONVIF: Capacidade de inclusão de novas câmeras ao sistema automaticamente pelo protocolo ONVIF.

RECURSOS PARA PROTEÇÃO DE DADOS:

33. Necessário compatibilidade com a LGPD (Lei Geral de Proteção de Dados) dispondo pelo menos das seguintes funcionalidades:

33.1. Deve possuir a opção de borrar rostos de pessoas que por circunstância compõe a imagem no ato do reconhecimento, mas que são alheias a pessoa de interesse cadastrada;

33.2. Deve possuir a opção de salvar detecções apenas das pessoas que estão cadastradas na base quando detectadas.

34. Deve possuir recurso para borrar objetos que não são parte do banco de dados de pessoas de interesse.

SEGURANÇA:

35. Deve possuir recurso para acesso a interface gráfica/sistema por meio de autenticação facial

36. Deve possuir recurso de monitoramento de sessão por re-autenticação facial do operador

37. Deve possuir recurso de identificação e histórico de usuário conectados, id de dispositivo único, endereço IP e se o usuário está online ou offline

38. Deve possuir recurso para bloquear o acesso de um determinado dispositivo diretamente na interface gráfica, bem como, desbloqueá-lo da lista de bloqueios

39. Deve possuir recurso para definir os tipos de arquivo que podem ser anexados ao cadastro das pessoas de interesse

40. Necessário trabalhar com vídeos de câmeras que estão conectadas à plataforma de reconhecimento facial (vídeo “on-line”) e também com vídeos que não estão integradas à plataforma de reconhecimento facial (vídeo “offline”) e neste caso, que estejam nos formatos de vídeo MP4, FLV, codec de vídeo H.264, sendo assim possível realizar buscas forenses, como por exemplo, vídeos oriundos de VMS, telefones celulares ou câmeras corporais (“body cam”).

41.41. Analítico Facial - Ter analíticos, que identificam:

41.1. Gênero

41.2. Barba

41.3. Idade

41.4. Uso de óculos de grau ou de sol.

41.5. Máscara de proteção respiratória posicionada corretamente e incorretamente

41.6. Emoções: bravo, medo, nojo, alegria, surpresa, tristeza e neutro.

42. Analítico Corporal

42.1. Ter analíticos, que identificam:

42.1.1. Cor da parte superior da roupa

42.1.2. Roupa superior (manga comprida, manga curta, sem manga)

42.1.3. Tipo de roupa (Jaqueta, agasalho, camisa, camiseta, vestido) 42.1.4. Cor da parte inferior da roupa

42.1.5. Tipo da parte inferior da roupa (calças, saia, bermuda)

42.1.6. Uso de chapéu, boné, bandana ou capuz

42.2. Ter analíticos que identificam fluxo de pessoas para um determinado período informando:

42.2.1. Número de visitantes

42.2.2. Quantos visitantes novos e quantos visitantes recorrentes 42.2.3. Idade média dos visitantes

42.2.4. Porcentagem distribuída entre Gêneros

43. Analítico Veicular - Ter analíticos, que identificam as seguintes características de um carro:

43.1. Tipo de carroceria

43.2. Fabricante

43.3. Modelo

43.4. Cor e placa de um carro

GERENCIAMENTO:

44. Deve permitir criar diferentes grupos de câmeras

45. Deve permitir criar diferentes listas de interesse (associada a cada câmera ou grupo de câmeras)

46. Deve permitir criar diferentes perfis de acesso/níveis de segurança, com pelo menos 3 níveis com a possibilidade de particularizar o que cada nível terá acesso

47. Ter a possibilidade de fazer filtragem simultânea dos seguintes fatores: por pessoa específica, de pessoas com e sem máscara de proteção respiratória, com uso da máscara de proteção respiratória imprópria, com e sem barba, com e sem óculos de grau, com e sem óculos de sol, por faixa de idade, por emoções, por câmera, por grupo de câmeras, por lista de interesse, por dia do evento.

INTEGRAÇÃO:

48. Disponibilizar API (“application programming interface”) aberta para integração com outros sistemas, possuindo diferentes métodos para chamada nos eventos de faces, corpos e carros

49. Possuir recurso de disparo de “Webhooks” para eventos relacionados a faces, corpos, carros e contadores

50. Possuir integração com VMS de mercado

51. Deve possuir recursos para receber, processar objetos em quadros (frames) e criar eventos a partir de um dispositivo de borda, por exemplo, controle de acesso.

LICENCIAMENTO:

52. Deve prever os seguintes tipos de licenciamento:

52.1. Baseado em validação “online”

52.2. Baseado em validação “offline” por assinatura digital de hardware (“assinatura do computador”).