

COGNEX VIDİ SUİTE : FAQ

Aspectos técnicos e de integração

Como funciona o Cognex ViDi Suite e quais métodos são utilizados?

Cognex ViDi Suite é uma implementação especializada de uma rede neural que foi otimizada para visão computacional sob condições fabris realistas. As especificações desta otimização são confidenciais, visto que são elas que tornam o Cognex ViDi Suite único.

O Cognex ViDi Suite é subjetivo ou objetivo?

O algoritmo de treinamento do ViDi introduz de maneira intencional uma pequena quantidade de aleatoriedade. Consequentemente, nós tipicamente medimos uma variação de 1-2% no resultado final mesmo utilizando o mesmo conjunto de dados. Para mais informações consultar a documentação técnica para mais detalhes (“By how much does the results accuracy vary from one training to another?”)

O quão rápido é o Cognex ViDi Suite?

O tempo de treino típico é de 2-5 minutos em um computador com uma placa GPU Nvidia de razoável alto desempenho. Nós providenciamos um *link* para uma tabela atualizada de GPUs disponíveis que incluem os possíveis desempenhos relativos para consulta (http://www.videocardbenchmark.net/high_end_gpus.html).

A velocidade de execução do aplicativo depende principalmente de três fatores:

- Tamanho da imagem: imagens maiores requerem mais tempo para serem processadas
- Tamanho do defeito: detecção de defeitos pequenos demoram mais do que defeitos grandes
- GPU: quanto mais GPUs (no máximo 4), e quanto mais novos ou potentes eles forem, maior será a velocidade.

Como uma estimativa, o tempo de execução levará entre 2ms (para um único caractere) e 2 segundos (para uma inspeção detalhada de um ambiente grande). Como regra geral, o cálculo computacional pode ser realizado a 10-15 Megapixels/segundo/GPU, ou 40-60 Megapixels/segundo para 4 GPUs de alto desempenho. Entretanto, testes deverão ser feitos para cada caso para obter uma estimativa mais precisa.

Qual é o tipo de dado que as ferramentas ViDi retornam como *feedback*?

Cada ferramenta providência dados XML com os seguintes resultados. Favor consultar a documentação do produto para detalhes da API.

ViDi Blue: coordenadas da imagem (X, Y), orientação dos atributos (θ), escala, pontuação, modelos geométricos, ângulos entre dois atributos.

ViDi Red Não Supervisionado: pontuação máxima, pontuação de cada pixel, centro de gravidade da área defeituosa, polígono delimitador, *hit map*, área de anomalia.

Vidi Red Supervisionado: pontuação máxima, pontuação de cada pixel, centro de gravidade da área defeituosa, polígono, *hit map*, área de anomalia.

Vidi Green: índice de confiança, 1ª melhor tipo de classe encontrada, 2ª melhor tipo de classe encontrada, etc.

A fonte da imagem é importante, ou é possível utilizar imagens de outra natureza com o Cognex ViDi Suite?

Cognex ViDi Suite tem suporte a imagens com extensão do tipo PNG, BMP, TIFF e JPEG, de 8 ou 16 bits e 1-4 canais. Contanto que seus dados estejam conforme essas especificações, as ferramentas podem processar suas imagens. Isto não implica nenhum desempenho específico; por exemplo, algumas fontes de dados podem conter mais ruídos que outras, ou estarem sujeitas a tendências que influenciam os algoritmos do ViDi. Porém, como regra geral, é possível utilizar imagens de fontes termográficas, raio-X e 3D com o ViDi Suite.

Qual é o maior tamanho de imagem capaz de ser processada adequadamente pelo ViDi?

Não existe um limite prático para treino/tempo de execução, sendo que o sistema pode facilmente manipular com imagens maiores que 1GB. No entanto, a GUI apenas é otimizada para imagens de até 10 MP. Imagens maiores podem causar inconsistências na experiência de interação do usuário com o programa, mesmo que o treino seja executado corretamente.

Existe algum tipo de HMI providenciado com o Cognex ViDi Suite ou ele é apenas uma biblioteca?

A interface de treino do Cognex ViDi Suite é um executável interativo *standalone* com sua própria UI, que produz seu próprio “*workspace*” (modelo treinado).

Em tempo de execução, os usuários programarão em uma API .NET para executar a ferramenta Red, Green ou Blue.

Existem três opções para exibir os resultados do ViDi.

- Usar a API ViDi para criar uma imagem com gráficos sobrepostos.
- Usar componentes ViDi WPF para desenhar gráficos em um *display* separado.
- Usar o dado resultante do ViDi para desenhar manualmente gráficos sobrepostos em um *display* VisionPro, que também permite exibir resultados da VisionPro.

No futuro, Cognex proverá métodos para combinar de maneira mais fluida os resultados da VisionPro e das ferramentas ViDi em um *display* comum.

A GPU de um sistema é exclusiva para a Cognex ViDi Suite ou ela também pode ser utilizados em outros aplicativos?

A GPU não é exclusivamente reservada para o Cognex ViDi Suite, porém, utilizá-la para outros aplicativos podem impactar o desempenho do ViDi.

Posso utilizar múltiplos GPUs?

Absolutamente. O Cognex ViDi Suite tem suporte a até 4 GPUs NVidia. O sistema irá detectar automaticamente e fazer uso destas GPUs para acelerar o aplicativo. Para mais detalhes, consultar no FAQ de Documentação Técnica o item “How to Build a 4-GPU Machine” e “By How Much will Processing Time Improve with Multiple GPUs”.

Como o ViDi pode se comunicar diretamente através de FFP ou I/O discretas?

ViDi é uma DLL com uma API .NET para as ferramentas Red, Green e Blue. Ela não inclui aquisição de imagens ou capacidade de I/O. Estas funções devem ser implementadas de maneira separada. Para simplificar isso, todas as novas remessas do ViDi incluirão uma cópia do VisionPro, que fornecerá aquisição de imagem e leitura/gravação de banco de dados, *display* de imagens, I/O e funções de comunicação industriais.

Capacidades de Aprendizagem/Treino.

Qual é o propósito do Development dongle?

O Development dongle providencia uma meio de treinar e qualificar o sistema, incluindo as ferramentas Red, Green e Blue. Como isto é por tempo limitado e não é adequado para implantação, o preço é inferior que o Development dongle de tempo de execução permanente.

Por quanto tempo é válido o Development dongle?

Um ano a partir do envio.

O que é necessário para treinar um arranhão/ranhura?

Como o Cognex ViDi Suite aprende por exemplos, é necessário apresentar dezenas de imagens que mostram a variedade de ranhuras que precisam ser detectados. Estes devem ser casos reais, exibindo variação em forma, tamanho, material, iluminação, etc – sejam quais forem as condições encontradas na linha de produção.

Como se organiza o Cognex ViDi Suite se, por exemplo, são inspecionados 20 vistas diferentes?

A resposta depende do quão consistentes são as imagens (e especialmente os defeitos). Se os defeitos são muito similares de vista para vista, um único treino pode ser suficiente. Se existe uma significativa variação no ambiente e na aparência dos defeitos (como campo de visão, iluminação, etc), então múltiplos *workspaces* precisarão ser criados. Não existe uma resposta simples que sirva para todos os casos.

Como se re-treina o Cognex ViDi Suite se surgirem novos defeitos?

Não deve haver necessidade de retreinar a ferramenta Red no modo Não Supervisionado. A ferramenta Red Não Supervisionado é treinada apenas nas

imagens corretas, que não apresentam defeitos, e portanto relatam quaisquer anomalias dessas imagens.

No modo Supervisionado, é desejado que imagens dos novos defeitos sejam adicionados ao conjunto de dados de treino. Os defeitos deverão ser destacados de maneira explícita, de forma que o sistema saiba o que constitui o novo tipo de defeito. Tenha em mente que a adição de novas imagens ao conjunto de dados de treino requer um re-treino completo. Não é possível adicionar novos dados sem um re-treino completo. Neste caso, Cognex recomenda revalidar o sistema no conjunto de dados de teste.

Com a visão tradicional eu sei exatamente o que é possível ou não. De que forma eu poderia saber como o Cognex ViDi Suite se comporta?

De fato, o ViDi não se comporta como uma ferramenta tradicional de processamento de imagens, com um comportamento definido de maneira explícita. Não há garantias que ele detectará todos os defeitos de um tipo, por exemplo. Mas em compensação, a vantagem do ViDi é sua capacidade de ser muito mais adaptável às condições do mundo real, e pode ser, de maneira mais fácil, ser implantado para resolver problemas que não podem ser resolvidos (seja de maneira parcial, ou total) com abordagens tradicionais da visão computacional.

Logo, é essencial começar com um conjunto de dados de treino representativo do problema que se deseja solucionar, categorizar as imagens e/ou os defeitos nela presentes de maneira correta, e cuidadosamente ajustar o sistema até atingir um comportamento esperado.

Como controlar as imagens de treino e achar o melhor conjunto de treino?

Por padrão, o sistema utiliza apenas 50% das imagens providenciadas para treino, selecionadas aleatoriamente. É possível alterar esta porcentagem para, por exemplo, incluir mais imagens no treino caso haja uma grande variação estética ou caso o conjunto de dados seja pequeno. É possível diminuir a porcentagem caso o conjunto de dados é muito grande e consistente. Também é possível selecionar e incluir imagens específicas, caso elas contenham atributos importantes que estão presentes apenas em poucas imagens.

Adicionalmente, existem alguns outros parâmetros que podem ser ajustados de acordo com os requerimentos de sua aplicação. O sistema também inclui um recurso de otimização interna (o "Parameter Search Tool") que irá automaticamente ajustar certos parâmetros, com base nas imagens e classificações. Isto auxilia na seleção de uma combinação ideal dos parâmetros.

Como é realizada a segmentação em uma aplicação de OCR?

Segmentação é realizada internamente pela ferramenta ViDi Blue, e não como uma etapa separada de pré-processamento anterior ao reconhecimento de caractere. É possível pensar na etapa de segmentação e reconhecimento como um único processo dentro da ferramenta Blue. Muito

do que é normalmente feito durante a segmentação é tratado em tempo de treino pela ferramenta Blue, onde ocorre o aprendizado da relação esperada entre os caracteres. Com impressão degradada (como em aplicações de jato de tinta em alta velocidade) a segmentação pode ser extremamente desafiadora usando técnicas tradicionais de OCR. Esta é um dos pontos fortes da ferramenta Blue.

Aplicações

Quais são as limitações do software ViDi?

ViDi Suite não é adequado para os seguintes tipos de aplicação:

- **Medições precisas** e metrologia
- **Classificação da qualidade de impressão.** ViDi não é um sistema baseado em fontes de caracteres, onde o usuário especifica o caractere esperado em cada posição. Em vez disso, ele é treinado no padrão geral, e às vezes pode se confundir caso um caractere mal impresso se assemelhe com um caractere válido.
- **Aplicações onde imagens em tempo de execução diferem significativamente** das imagens de treino em relação a luminosidade, contraste, FoV, etc., mas ainda são consideradas aceitáveis.
- **Não pode ser treinado a partir de dados CAD**

Quais tipos de aplicações são solucionáveis com o ViDi?

Conext ViDi Suite é especialmente apropriado para três tipos de aplicações:

- **Inspeção cosmética** como arranhões, manchas, lascas e ausência ou excesso de tinta, onde soluções tradicionais baseados em regras são ineficientes ou extremamente complexas.
- **Localização de atributos** em partes deformadas ou inconsistentes onde PatMax não é um candidato viável. Note que a informação da posição resultante não é tão precisa quanto PatMax, porém, ainda assim é possível encontrar com maior eficiência objetos cuja aparência seja menos consistente.
- **Reconhecimento óptico de caracteres em textos deformados** como em borracha, metal estampado ou proveniente de jatos de tinta de alta velocidade.
- **Classificação de defeitos**
- **Classificação de produtos**

Onde está a fronteira entre Cognex ViDi Suite e visão tradicional?

Algumas aplicações são claramente mais adequadas para visão computacional tradicional, como alinhamento de precisão, metrologia e leitura de código 2D. Outras são melhores para o Cognex ViDi Suite, como inspeção cosmética ou OCR de caracteres deformados. A fronteira inclui muitas aplicações de localizar atributos e OCR, ou inspeções simples/consistentes.

Nestes casos, nossa recomendação é testar ambas as abordagens e verificar qual obtém o melhor desempenho. Outras considerações incluem o custo de

acelerar o *hardware* utilizando GPU e a facilidade/complexidade de treinar o sistema. Apesar disto, alcançar um resultado preciso e estável de visão computacional deve ser a principal preocupação.

O que é necessário para realizar um teste de viabilidade técnica?

É necessário coletar um conjunto de dados de imagens corretas e defeituosas sob as condições esperadas de luminosidade, assim como entender o critério do defeito do cliente.

- Para detecção de objetos, recomendamos 30-50 amostras de imagens representando a variedade de casos esperados.
- Para OCR, 30-50 amostras de cada caractere, amostradas na variação de cada posição possível.
- Para Inspeção Não Supervisionada, 30-50 imagens corretas e 30-50 imagens defeituosas.
- Para Inspeção Supervisionada, 30-50 imagens de cada tipo de defeito ou atributo que deve ser segmentado.

Tenha em mente que o número de imagens necessárias está intimamente relacionada com a precisão que o cliente procura. Por exemplo, caso seja necessário 99.9% de precisão, pode ser necessário 1000 imagens, enquanto que 99% de precisão pode precisar de apenas 100 imagens.

Quanto *pixels* um “defeito” precisa conter para ser detectável no ViDi?

O *software* ViDi é capaz de detectar defeitos tão pequenos quanto um único *pixel*. No entanto, “defeitos” deste tamanho podem ser consequência de ruído da câmera ou da luminosidade. O tamanho mínimo do defeito deve ser configurado tendo como base a qualidade da imagem disponível e do tamanho real confiável do defeito que precise ser detectado.