USO DO MÉTODO LASSO PARA RESOLUÇÃO DO PROBLEMA DE RECONHECIMENTO FACIAL

Pedro Henrique Matos de Lima (PIC/UEM), Ronaldo Lopes (Orientador), Francisco Nogueira Calmon Sobral (Coorientador). E-mail: rlopes2@uem.br

Universidade Estadual de Maringá, Centro de Ciências Exatas, Maringá, PR.

Área e subárea do conhecimento: Matemática/Matemática Aplicada

Palavras-chave: método de descenso coordenado; representação esparsa; classificação supervisionada.

RESUMO

Neste trabalho estudamos um método de descenso coordenado para resolver o problema de otimização conhecido como LASSO e utilizamos esta classe de problemas para resolver o problema de reconhecimento facial. O problema de reconhecimento facial está relacionado a automatização da identificação da imagem do rosto de um indivíduo dentro de um banco de dados contendo imagens das faces de um grupo prefixado. Selecionados um banco de dados bem estabelecido na literatura, testamos formas de filtro dos pixels das imagens e parâmetros para efetuar o reconhecimento facial por meio da resolução do problema LASSO, aplicando estratégias de identificação e validação propostas em literaturas do gênero. Discutimos o desempenho tanto da identificação quanto da validação das faces dos indivíduos, escolhendo uma fração dos pixels de cada imagem de forma sequencial e aleatória. Para o banco de dados analisado, observamos que a escolha aleatória de pixels foi mais eficiente na identificação das faces e que um grande número de parâmetros do modelo permitiu taxas de acerto superiores a 90%. A validação, por sua vez, apresentou maior sensibilidade tanto aos parâmetros do modelo quanto ao *threshold* selecionado, evidenciando que a eficácia deste processo é mais dependente dessas escolhas.

INTRODUÇÃO

Com a crescente evolução tecnológica, a automatização de processos vem sendo uma demanda cada vez mais requisitada pelos processos produtivos. Em computação científica, buscamos essa automatização, em muitas situações, por meio da resolução de problemas de otimização. O problema de reconhecimento facial acompanha esta dinâmica.

No artigo de Wright *et al.* (2009), os autores propuseram a resolução de um problema de otimização conhecido como *Basis Pursuit*, visando construir uma solução que promova um equilíbrio entre a solução do problema de quadrados mínimos juntamente com várias coordenadas nulas, garantindo um bom desempenho empírico no problema de reconhecimento facial supervisionado.

O problema *Basis Pursuit* pode ser reescrito como um problema quadrático com restrições, ao custo de dobrar o número de variáveis, e, no artigo supracitado, foi utilizado uma técnica de otimização chamada Programação Quadrática Sequencial para resolver os problemas de otimização abordados. Existe um famoso problema de otimização equivalente ao problema *Basis Pursuit* chamado de LASSO (*Least Absolute Shrinkage and Selection Operator*), introduzido em Tibshirani (1996), este com técnicas de descenso coordenado bem estabelecidas e eficientes para encontrar uma solução para ele.
Este trabalho visou estudar técnicas de descenso coordenado, presente em Lopes (2018), para resolução do problema LASSO e aplicar essas técnicas ao problema de reconhecimento facial, desenvolvendo habilidades de programação, estudo de técnicas de otimização e investigação científica na escrita, descrição e análise dos resultados obtidos.

MATERIAIS E MÉTODOS

Para desenvolvimento dos métodos e manipulação dos dados, escolhemos utilizar a linguagem de programação Julia, uma linguagem de programação científica que combina ferramentas e facilidade na manipulação de objetos inspiradas em várias linguagens como: Python, MATLAB. Implementamos o método de descenso coordenado e as técnicas de identificação e validação para o reconhecimento facial nesta linguagem.
Utilizamos um banco de imagens dividido entre 28 indivíduos, conhecido como *Extended Yale Face Database B*, o qual contém 585 imagens faciais em escala de cinza para cada indivíduo, com resolução de 640x480 pixels, capturadas sob diferentes condições de iluminação. Para diminuir o volume de dados manipulados, retirar regiões irrelevantes da imagem e se adequar a teoria de *compressed sensing,* inicialmente fizemos um corte retangular, reduzindo a dimensão das imagens para 440×380 pixels, para evitar que os testes abordassem regiões indesejadas. Com o intuito de empregar as imagens a técnica de reconhecimento facial, cada imagem foi transformada em um vetor, e em seguida, agrupados como colunas de uma matriz.
Na sequência, empregamos outra estratégia para redução de dimensionalidade, optamos por executar os algoritmos com uma fração dos dados já reduzidos pelo corte retangular, escolhemos dois critérios de seleção: aleatório (1%, 2%, 4% e 8% dos pixels) e sequencial (1%, 4% e 16% dos pixels), permitindo comparar o desempenho do método sob diferentes configurações. A identificação das imagens foi feita variando a escolha de parâmetros do problema LASSO e alterando a escolha da imagem do banco de dados que tentaria ser identificada.
Adicionalmente, investigamos a capacidade de validação de faces por meio da representação esparsa. Para isso, acrescentamos ao conjunto de dados 10 imagens, sem relação com as faces previamente consideradas, e aplicamos a elas as mesmas técnicas de redução de dimensionalidade. Os experimentos variaram os parâmetros do modelo e as imagens submetidas à validação, com o objetivo de

analisar possíveis valores de *threshold* capazes de distinguir adequadamente imagens pertencentes ao conjunto de faces daquelas externas a ele.

RESULTADOS E DISCUSSÃO

Aplicamos o algoritmo denominado SRC (*Sparse Representation-based Classification*), que é responsável por classificar a imagem de entrada como o indivíduo do banco de imagens de treinamento cuja classe apresenta o menor resíduo de reconstrução. Para aplicação dele, é requerido a resolução de um problema LASSO, cuja formulação depende do parâmetro de regularização λ, responsável por equilibrar a qualidade da reconstrução da imagem de teste e a esparsidade da solução obtida. Foram usados nos testes seis valores diferentes: $\lambda = 10^{n} \times \|A^{T}b\|_{\infty}$, com $n \in \{-1,-2,-3,-4,-5,-6\}$.

A seleção aleatória de pixels manteve taxa de reconhecimento de 100% de acerto em todos os experimentos, independentemente de λ ou da porcentagem de pixels utilizada, evidenciando que informações distribuídas pela imagem preservam as características discriminantes necessárias. Já a seleção sequencial teve desempenho inferior com poucos pixels, com taxas de 60% a 90% de acerto com 1% dos pixels, melhorando progressivamente até atingir 100% com 4% ou 16% dos pixels. Calculando uma média da precisão de acerto de cada valor de λ para todas as situações de escolha de pixels sequencial e aleatória, observou-se que os melhores desempenhos foram atingidos para $n \in \{-3,-4,-6\}$, com taxa média de 96,7% de precisão.

Além do reconhecimento facial, é igualmente importante determinar se uma imagem de teste pertence ao conjunto de indivíduos do banco de dados. Para isso, utilizamos o SCI (*Sparse Concentration Index*), um índice que mede o grau de concentração dos coeficientes da solução em uma única classe, resultando sempre em um valor dentro do intervalo [0,1]. Se a solução está completamente concentrada em uma única classe, SCI=1. Se os coeficientes da solução estão igualmente distribuídos entre todas as classes, em relação a norma da soma, SCI=0. Assim, utilizamos um limiar $\tau \in (0,1)$ para validar se a imagem pertence ao banco ou não.

Os testes para avaliar a escolha de um bom limiar τ foram conduzidos da seguinte forma: para cada um dos 6 valores de λ, mesmos da etapa de classificação, realizamos 5 testes com uma face de um indivíduo e 5 testes com uma das 10 imagens genéricas, ambos escolhidos aleatoriamente por meio de uma distribuição uniforme. Em seguida, resolvemos o problema LASSO para diferentes valores λ e a partir da solução aproximada, calculamos o índice SCI correspondente, a fim de identificar boas escolhas para o *threshold*. Analisamos os valores mínimos, máximos e médios do SCI obtidos tanto para faces de indivíduos quanto para imagens genéricas.

Com base nos testes, recomenda-se utilizar $\lambda = 10^{-3} \times \|A^{T}b\|_{\infty}$, combinado a um limiar de *threshold* $\tau \in (0{,}1;0{,}3)$, para indicar se uma imagem de teste é válida ou não em

relação a um conjunto de faces prefixado. Essa faixa de parâmetros proporciona uma margem de segurança maior, aumentando a chance de validação correta pela técnica proposta.

CONCLUSÕES

Neste trabalho, foi investigado o reconhecimento facial por meio de representação esparsa, utilizando o problema LASSO e, como estratégia de resolução, o método de descenso coordenado, com atualizações baseadas no operador *Soft-Thresholding*. Os resultados experimentais demonstraram um desempenho satisfatório no reconhecimento facial mesmo com o uso de uma pequena fração dos pixels das imagens, com a seleção aleatória de pixels apresentando desempenho superior à seleção sequencial, para o banco de dados analisado. Esse resultado indica que informações distribuídas espacialmente na imagem contribuem de forma significativa para o reconhecimento. O *Sparse Concentration Index* (SCI) mostrou-se igualmente relevante, complementando a classificação ao fornecer um critério para validar se as imagens analisadas pertencem ou não ao banco de imagens. De modo geral, a combinação entre LASSO, representação esparsa e SCI mostrou-se uma alternativa eficiente e robusta para o reconhecimento e a validação facial.

REFERÊNCIAS

TIBSHIRANI, R. Regression Shrinkage and Selection via the Lasso. **Journal of the Royal Statistical Society. Series B (Methodological)**, 58(1), 267–288, 1996.

WRIGHT, J. et al**.** Robust face recognition via sparse representation. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, [s.l.], v. 31, n. 2, p. 210-227, 2009.

LOPES, R. **Métodos de descenso coordenado por blocos e identificação das restrições ativas em otimização de porte enorme.** Tese de Doutorado – Instituto de Matemática e Computação Científica, Universidade Estadual de Campinas, Campinas, 2018.