

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2022年7月7日(07.07.2022)



(10) 国際公開番号

WO 2022/145343 A1

(51) 国際特許分類:
G06V 30/194 (2022.01) G06V 30/412 (2022.01)
G06V 30/14 (2022.01)

(21) 国際出願番号: PCT/JP2021/047935

(22) 国際出願日: 2021年12月23日(23.12.2021)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願 2020-219612 2020年12月28日(28.12.2020) JP

(71) 出願人: 有限責任監査法人トーマツ
(DELOITTE TOUCHE TOHMATSU LLC) [JP/
JP]; 〒1008360 東京都千代田区丸の内三丁目2
番3号丸の内二重橋ビルディング Tokyo (JP).

(72) 発明者: シェイク ホサイン シャハリアル(Sheikh,
Hossain Shariar); 〒1008360 東京都千代田区丸

の内三丁目2番3号丸の内二重橋ビルディング
有限責任監査法人トーマツ内 Tokyo (JP).

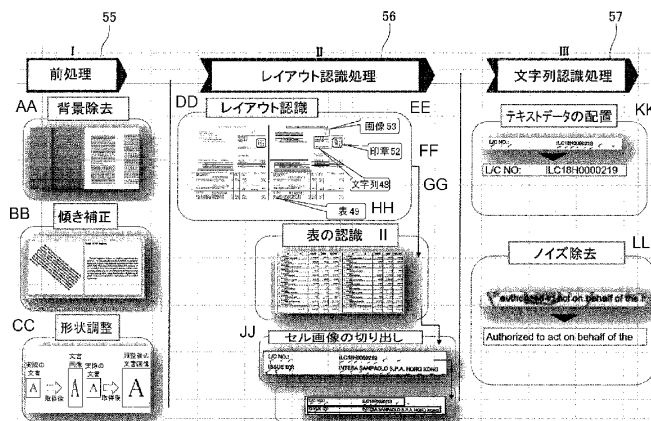
(74) 代理人: 特許業務法人白坂 (SHIRASAKA &
PATENT PARTNERS); 〒1000004 東京都千
代田区大手町一丁目6番1号 大手
町ビル4階 Tokyo (JP).

(81) 指定国(表示のない限り、全ての種類の国内保
護が可能): AE, AG, AL, AM, AO, AT, AU, AZ,
BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH,
CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, KE, KG, KH,
KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,
MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ,
NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT,
QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) Title: ARCHITECTURE FOR DIGITALIZING DOCUMENTS USING MULTI-MODEL DEEP LEARNING,
AND DOCUMENT IMAGE PROCESSING PROGRAM

(54) 発明の名称: 多モデル深層学習による文書のデジタル化アーキテクチャ、文書画像処理プログラム

[図3]



- 55 Preprocessing
- 56 Layout recognition processing
- 57 Character string recognition processing
- AA Background elimination
- BB Tilt correction
- CC Shape adjustment
- DD Layout recognition
- EE Image 53
- FF Seal 52
- GG Character string 48
- HH Table 49
- II Table recognition
- JJ Cell image segmentation
- KK Text data arrangement
- LL Noise elimination

(57) Abstract: The purpose of the present invention is to convert a character string contained in a document image to text data by a method different from conventional optical character recognition. In the present invention, an electronic document generation device is provided with: a document image acquisition unit that acquires a document image obtained by imaging a document; a character string recognition unit that recognizes a character string contained in the document image acquired by the document image acquisition unit using a character string learning model that has learned correspondence between document images and character strings contained in the document images and outputs text data of the recognized character string; and an output unit that outputs the text data as text for an electronic medium.

[続葉有]

ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, WS, ZA, ZM, ZW.

- (84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告(条約第21条(3))

(57) 要約: 文書画像に含まれる文字列を、従来の光学文字認識とは異なる手法によって、テキストデータに変換することを目的とする。電子文書生成装置は、文書を画像化した文書画像を取得する文書画像取得部と、文書画像と当該文書画像に含まれる文字列との対応関係を学習した文字列学習モデルを用いて、文書画像取得部に取得された文書画像に含まれる文字列を文字認識し、当該文字列に係るテキストデータを出力する文字列認識部と、テキストデータを電子媒体のテキストとして出力する出力部とを備える。