



2021년 12월 7일 수신, 2021년 12월 22일 수락, 발행일 2021년 12월 23일, 현재 버전 날짜 2021년 12월 31일.

디지털 객체 식별자 10.1109/ACCESS.2021.3138240

# 점자를 위한 딥 러닝 전략 문자 인식

TASLEEM KAUSAR AND M.  사자드 만주르<sup>1</sup>, 아디바 카우사르<sup>2</sup>, 윤 루 ,  
Institute of Technology, Mardan, Pakistan 1Department of Computer Science and Information Technology, University  
of Narowal, Narowal, Punjab 51600, Pakistan 3School of Narowal Computer Science and Engineering, Huizhou  
University, Huizhou, Guangdong 516007, China 4Department of Electrical Engineering, University of Gujrat, Gujrat,  
Punjab 50700, Pakistan 교신저자: Yun Lu (luyun\_hit@163.com)

이 작업은 그랜트 2020A151110498에 따라 광동성의 기초 및 응용 기초 연구 기금의 공동 기금에 의해 부분적으로 지원되었으며, 부분적으로는 그랜트 2020JB058에 따라 후이저우 대학의 교수 및 박사 과학 연구 재단에 의해 부분적으로는 향상된 독립 계획 프로젝트에 의해 지원되었습니다. Grant hzu202018에 따른 Huizhou University의 혁신 능력, 부분적으로 Grant 62176102에 따라 중국 국립 자연 과학 재단에 의해.


● **개요** 시각 장애가 있는 사람들은 읽기, 쓰기 및 의사소통을 위해 점자 언어를 사용합니다.

점자 언어의 기본 구조는 3행 2열 셀로 배열된 6개의 점으로 구성되며 시각 장애인이 손가락 터치를 통해 식별할 수 있습니다. 그러나 점자 문자를 형성하는 점의 패턴을 기억하는 것은 어렵습니다. 이 연구는 자동 점자 문자 인식을 위한 새로운 접근 방식을 제시합니다. 설계된 접근 방식은 두 가지 주요 단계에서 작동합니다. 첫 번째 단계에서는 여러 이미지 전처리 기술을 사용하여 이미지 정렬 및 향상이 수행됩니다. 두 번째 단계에서는 제안된 경량 합성곱 신경망(CNN)을 사용하여 문자 인식을 수행합니다. CNN은 광학 문자의 정확한 인식에 대한 약속을 보여줍니다. 따라서 우리는 점자 문자 인식을 위해 최근에 제안된 몇 가지 최첨단 CNN 네트워크를 채택했습니다. 네트워크를 가볍게 만들고 인식 성능을 향상시키기 위해 원본 CNN의 일부 모듈을 계산 비용이 적은 IRB(inverted residual block) 모듈로 교체하는 전략을 제안했습니다. 이 작업의 참신함은 CNN 모델 설계 및 출력 성능에 있습니다. 시각 장애인에게서 얻은 두 가지 공개 벤치마크 점자 데이터 세트에 대한 실험을 통해 설계된 설정의 효율성을 실행했습니다. 영어 점자 및 중국어 양면 점자 이미지(DSBI) 데이터 세트에서 제안 모델은 각각 95.2% 및 98.3%의 예측 정확도를 나타냅니다. 모델의 보고된 테스트 시간은 영어의 경우 약 0.01초, DSBI 점자 이미지의 경우 0.03초입니다. 최신 기술에 비해 설계된 방법은 강력하고 효과적이며 점자 문자를 효율적으로 식별할 수 있습니다. 향후에는 접근 가능한 사용자 인터페이스를 통해 제안된 점자 인식 방식의 기능적 성능을 테스트할 예정입니다.

● **색인 용어** 점자 이미지, 이미지 정렬, 주성분 분석, Wiener 필터링, 컨볼루션 신경망, 역 잔차 블록.

## I. 서론

세계보건기구(WHO)에 따르면 전 세계적으로 약 22억 명이 실명 또는 시각 장애인이라고 합니다[1]. 이러한 시각 장애인들은 글을 읽고 쓰는 것이 어렵기 때문에 손가락으로 만지는 감각으로 읽을 수 있는 돌출된 점 체계인 점자를 사용합니다[2]. 점자 시스템의 기본 구조는 그림 1a와 같이 3×2 순서로 정렬된 6개의 점으로 구성된 매트릭스입니다. 점자 셀의 각 문자는

이 원고의 검토를 조정하는 부편집장과  
출판을 승인한 사람은 Shovan Barma였습니다. .

이 6개의 점을 특수한 방식으로 배열하여 형성됩니다.

따라서 6개의 위치에 대해 임의의 조합으로 점을 올릴 수 있으므로 총 64개의 조합이 가능합니다 ( $2^6 = 64$ ). 점자에서 모든 문자는 셀에 융기된 점으로 형성된 패턴으로 식별됩니다. 점자 점의 다양한 조합을 통해 형성된 점자 문자, 알파벳 및 기호의 코드는 그림 1b와 같습니다.

그 효과로 인해 점자 시스템은 서면 의사 소통을 위해 시각 장애인이 전 세계적으로 사용합니다. 그러나 많은 사람들, 특히 평범한 사람들은 점자 문자를 식별하지 못합니다.

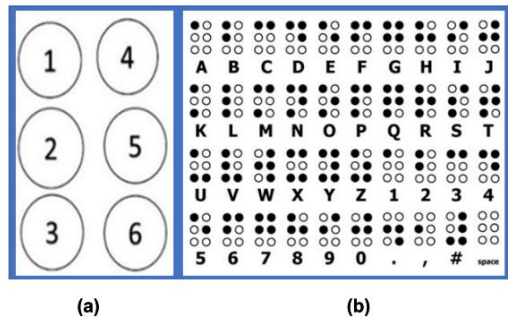


그림 1. (a) 점자점 행렬 (b) 점자점 조합을 사용한 알파벳.

최근 딥 러닝은 다양한 영역에서 성공적으로 주요 발전을 이루었습니다[3]-[6]. 심층 학습 방법, 특히 컨볼루션 신경망(CNN)은 자연 이미지 분류[7]-[9], 물체 감지[10]-[12], 이미지 분할[13]에서 최첨단 기술을 극적으로 향상시켰습니다. [14]-[15], 음성 인식[16], [17], 의료 이미지 분석[18]-[21]. 딥 러닝 기술은 이미지 인식에서 엄청난 발전을 이루었고 고성능 결과를 달성했지만 딥 러닝은 점자 문자 인식에 거의 사용되지 않습니다.

이 현대 시대에는 점자 문자 인식을 위한 방법 탐색에 대한 관심도 증가하고 있습니다[22]-[24].

그러나 점자 인식을 위한 심층 모델 사용에 대한 몇 가지 논문이 있습니다. 시각 장애인과 일반인의 차이를 줄이고 시각 장애인이 점자 언어를 빠르게 읽을 수 있도록 하기 위해 자동 점자 인식 작업이 수행됩니다[25]-[28]. 또한 [29]에서 저자는 CNN(Convolutional Neural Network)과 RCSA(Radio Character Segmentation Algorithm)를 사용하여 점자 문자를 인식하고 이를 영어로 변환했습니다.

그들은 제안한 방법의 높은 정확도를 주장했습니다.

94.42%의 정확도를 얻은 한자 인식 모델이 [30]에서 제안되었다. [30]에서 제안한 방법은 자연어에서 문자를 읽을 수 있도록 확장되었습니다.

점자 문자를 힌디어로 변환하기 위해 [31]에서는 기존의 시퀀스 매핑 방법과 딥 러닝 방법을 조합하여 출력을 음성으로 생성했습니다. [32]에서 저자는 실시간 점자 문자 인식을 위해 YOLO를 사용했습니다. [33]에서는 N-Best 알고리즘을 사용하여 만다린 점자 단어를 한자로 번역하는 소위 새로운 기법을 제안했습니다. 저자는 제안된 기법이 전체적으로 94.38%의 번역 정밀도를 얻었다고 주장했다.

[34]에서 저자는 점자 음악 이미지에서 점자 음악을 인식하기 위한 딥러닝 전략을 사용했습니다. 모델 성능을 높이기 위해 이미지 전처리 작업을 거의 적용하지 않았습니다. 시각 장애인이 점자 문자를 읽을 수 있도록 돕기 위해 [35]에서 또 다른 딥 러닝 기술이 사용됩니다. 그들은 Naive Bayes(NB), Decision Trees(DT), SVM 및 KNN과 같은 고전적인 기계 학습 기술과 결과를 비교했습니다.

서로 다른 알고리즘을 공정하게 비교할 수 있는 오픈 소스 데이터 세트가 거의 없습니다. 공개적으로 이용 가능한 데이터 세트의 부족은 또한 인식 알고리즘의 정확성을 평가하기 위한 심층 신경망의 훈련을 방해합니다. 위에서 언급한 작업들에서 제공하는 성능 값은 독자적인 데이터셋으로 측정된 것이기 때문에 점자 이미지 인식에서는 제안된 작업과 이전 작업의 공정한 비교가 불가능하다. Li et al. [36] BraUNet이라는 깊은 의미론적 분할 프레임워크를 제안했습니다. BraUNet은 인식 성능을 향상시키기 위해 보조 학습 전략과 후처리 전략을 사용하는 표준 UNet 모델[37]을 기반으로 합니다. Li et al. [38]은 또한 전통적인 기계 학습 계단식 분류기를 사용하는 TS-OBRI라는 또 다른 점자 인식 기술을 제안했습니다. [28]에서 저자는 광학 점자 문자 인식을 위해 RetinaNet[39]이라는 객체 감지 네트워크를 사용했습니다. U-Net, BraUNet, TS-OBRI 및 RetinaNet의 모든 모델은 DSBi 데이터 세트 [40]에서만 학습됩니다. DSBi 데이터 세트[40]는 레이블이 지정된 점자 텍스트 이미지가 있는 공개적으로 사용 가능한 유일한 데이터 세트이며 작성자에게 깊은 감사를 표하고 싶습니다.

공개적으로 사용 가능한 두 개의 영어 점자 데이터 세트[41]와 중국어 양면 점자 이미지 데이터 세트[40]에서 모델 성능을 평가했습니다.

초기 심층 CNN 모델, 예를 들어 AlexNet[42] 및 VGGNet[43]은 주로 입력 이미지의 해상도에 대한 엄격한 규칙이 있는 컨볼루션 레이어와 완전히 연결된 레이어로 구성됩니다. 이는 이러한 네트워크가 주로 두 가지 유형의 계층, 즉 컨볼루션 계층과 완전 연결 계층으로 구성되기 때문입니다. 컨볼루션 계층은 임의 크기의 입력 이미지를 받아 모든 크기의 출력 피쳐 맵을 생성할 수 있습니다. 컨볼루션 레이어의 출력 특징 맵은 완전 연결 레이어의 입력으로 사용됩니다. 본질적으로 완전 연결 계층은 고정된 크기의 입력을 받아들이므로 고정 크기 제약 조건은 완전 연결 계층에 의해 부과됩니다. 현재 CNN(예: Inception[44], ResNet[45] 및 DenseNet[46])은 전체 컨볼루션 계층을 포함하지만 글로벌 평균 풀링(GAP) 계층[47]도 채택합니다.

이 계층은 많은 수의 매개변수 문제를 해결하므로 네트워크 계산이 상당히 줄어듭니다.

한편, GAP 레이어는 모든 크기의 이미지 입력을 구현하기 위해 높이와 너비 차원의 평균을 취함으로써 출력 기능 맵을 압축했습니다. 최근 InceptionV3[44]라는 또 다른 모델이 작은 컨볼루션 커널을 사용하여 설계되었습니다. 이 모델에서 큰 컨볼루션 커널은 불필요한 레이어를 제거하여 깊이와 너비가 증가한 작은 커널로 분해됩니다. 이러한 수정은 InceptionV3의 매개변수와 계산 비용을 크게 줄입니다. 한편, 네트워크 깊이와 너비의 증가는 주로 네트워크의 기존 컨볼루션 계층을 개선합니다. DenseNet[46]은 각 레이어의 입력이 이전 레이어의 출력을 사용하는 밀집 연결을 사용했습니다. 이러한 조밀한 연결 패턴은 기울기 소실 문제를 완화합니다. 결과적으로 출력 기능이 강화되어 더 적은 매개변수로 더 나은 성능을 이끌어냅니다.

DenseNet 네트워크의 장점을 바탕으로 우리는

최종 제안된 심층 CNN 모델의 백본 네트워크로 사용됩니다.

또한 점자 이미지에서 문자를 자동으로 인식하는 데는 몇 가지 제한 사항이 있습니다. 휴대용 카메라 장치 및 스캐너로 캡처한 이미지는 일반적으로 대비가 낮고, 시각적으로 방해받고, 정렬이 어긋나고, 조명 변화가 있고, 불규칙한 노이즈로 오염되어 있습니다. 이러한 입력 이미지 왜곡은 심층 CNN 모델의 감지 성능에 영향을 미치고 점자 문자의 자동 인식을 어렵게 만들 수 있습니다. 따라서 CNN 모델에 공급하기 전에 표준 방향으로 처리하고 정렬해야 합니다. 심층 모델이 이미지를 올바르게 인식할 수 있도록 이미지 전처리가 중요합니다. 또한 심층 CNN에서 더 나은 정규화를 얻을 수 있는 유망한 솔루션입니다.

문헌에서는 이미지 필터링, 히스토그램 균등화[48], 색채 적응[49]과 같은 여러 이미지 전처리 기술이 이미지 선명도를 높이기 위해 적용되었습니다.

[50]의 저자들은 사람의 얼굴, 지문, 홍채 생체 특징을 이용한 보안 시스템 개발에 사용되는 다양한 이미지 전처리 방법에 대해 논의했습니다. 마찬가지로 이미지 정렬에 많은 작업이 수행되었습니다.

[51]에서 저자는 주성분 분석(PCA) 기반 이미지 정렬 알고리즘을 제안했습니다. [52]에서는 라인 가이드 로컬 워핑을 이용한 이미지 스티칭을 위한 2단계 이미지 정렬 방식이 제안되었다. [53]에서 관절면 정렬을 위한 방법이 제안되었고 기존 기술과 비교되었다. 그들은 관절 이미지 정렬 과정에서 얼굴의 경직된 변형과 경직되지 않은 왜곡을 고려했습니다.

[54]에서 Likassa는 아핀 변환과 Frobenius & norms를 사용하여 선형 상관 데이터 세트에 대한 이상치와 심한 희소 노이즈의 영향을 처리하기 위한 새로운 이미지 정렬 알고리즘을 제안했습니다. 최근 Huang et al. [55]는 전형적인 낮은 순위 쌍 정렬 이중 선형 분류 네트워크(LRPABN)를 제안했습니다. 설계된 모델에서 그들은 이미지 특징을 쿼리 특징과 일치시키기 위해 특징 정렬 레이어를 사용했습니다. [56]에서는 장면에 배치된 기준 마커를 사용하여 오정렬로 인한 시각적 아티팩트를 줄이는 최적화 프레임워크를 제안합니다.

이미지 인식에서 심층 CNN을 사용하는 데는 두 가지 주요 과제가 있습니다. (i) 정확도를 어떻게 더 높일 수 있습니까? (ii) 네트워크 계산 비용을 어떻게 줄일 수 있습니까? 본 논문에서는 점자 인식 문제를 이미지 분류 문제로 다루어 보다 효과적이고 적합한 CNN 기반 점자 인식 알고리즘을 구축하였다. 이전에 제안된 심층 CNN의 높은 계산 복잡성은 실시간 애플리케이션에 바람직하지 않았습니다. 기존의 접근 방식과 달리 네트워크 계산 비용 문제에 집중하여 가장 적합한 이미지 전처리 기법과 경량의 심층 컨볼루션 신경망을 사용하는 점자 문자 인식을 위한 새로운 접근 방식을 제안했습니다. 네트워크 계산 비용 문제를 해결하기 위해 역전차 블록(IRB) 모듈[57]을 사용한 동적 유형 CNN 수정 기법이 원본 CNN 모델에 사용되었습니다. 우리는 다음을 사용하여 제안한 방법의 성능을 테스트했습니다.

수정된 모델의 백본으로 널리 사용되는 다양한 최첨단 네트워크[43]–[44]. 우리가 아는 한, 경량 심층 CNN 모델을 사용하여 점자 문자 인식을 성공적으로 수행한 것은 우리 작업이 처음입니다. 실험 분석 결과 제안하는 기법이 인식 정확도를 향상시킬 뿐만 아니라 모델 파라미터와 계산 비용을 감소시키는 것으로 나타났다. 우리 작업의 주요 기여는 다음과 같이 논의됩니다.

- 자동 점자 문자 인식을 위한 새로운 접근 방식을 제안합니다. 제안된 접근 방식은 이미지 전처리 기술과 고급 경량 컨볼루션 신경망 모델의 조합을 기반으로 합니다.
- 빠른 경량 CNN 네트워크는 계산 비용이 적은 IRB 모듈을 사용하여 설계되었습니다. 네트워크 효율성을 향상시키기 위해 다양한 효과적인 전략도 적용됩니다.
- 제안된 네트워크는 두 가지 유형의 점자 데이터 세트에서 효과적으로 작동하며 예측 정확도 및 인식 속도 측면에서 최신 기술을 능가합니다.

English Braille 및 DSBi 데이터 세트에서 제안된 접근 방식은 각각 95.2% 및 98.3%의 예측 정확도를 나타냅니다. 영어와 중국어 데이터 세트에 대해 제안된 모델의 보고된 테스트 시간은 각각 약 0.01초와 0.03초입니다.

이 논문은 다음과 같이 더 구성되어 있습니다. 섹션 II에서는 이미지 전처리 단계에 대해 설명합니다. II장에서는 제안하는 방법에 대한 기술적인 설명을 한다. 실험 설정 및 결과는 섹션 III에 설명되어 있습니다. IV장에서는 논문을 마무리한다.

## II. 방법론 A. 전처리

이 점자 문자 인식 방식은 크게 두 단계로 구성된다. 첫 번째 단계에서는 일련의 이미지 전처리 기법을 사용하여 이미지 정렬 및 향상을 수행하고 다음 단계에서는 제안된 컨볼루션 신경망을 사용하여 문자 인식을 수행합니다. 점자 문자 인식의 흐름도는 그림 2에 나와 있습니다. 이미지 정렬 및 향상에 사용되는 전처리 기술은 다음 섹션에서 설명합니다.

### 1) 주성분과의 이미지 정렬 분석

스캐너 및 카메라 장치로 얻은 점자 이미지는 일반적으로 기하학적으로 교란됩니다. 점자 이미지는 시각 장애인이 손가락 터치를 사용하여 식별할 수 있는 3×2 매트릭스로 정렬된 6개의 점으로 구성됩니다.

이미지가 정확한 위치에 있지 않으면 시각 장애인이 캐릭터를 예측하는 데 문제가 발생합니다. 또한, 기하학적으로 교란된 이미지에 대해 학습된 자동 점자 문자 인식 시스템의 성능도 이미지 회전 문제에 의해 현저하게 영향을 받습니다. 이러한 이미지를 올바르게 인식하기 위해서는 기하학적 왜곡을 해결하고 이미지를 정렬하는 것이 중요합니다.

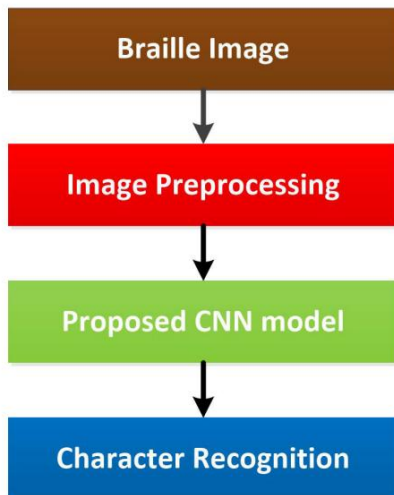


그림 2. 점자 문자 인식 프로세스의 흐름도.

적절한 위치에. 본 연구에서는 주성분 분석 알고리즘[51]을 사용하여 이미지 정렬을 달성했습니다.

PCA [51]를 사용한 이미지 정렬은 다음 단계로 설명됩니다. 단계 1: 입력 이미지가 주어지면 값이 1인 픽셀의 좌표(열 및 행)가 단일 행렬  $Z$ 에서 추출됩니다.

같이:

$$Z = \begin{bmatrix} c_1 & r_1 \\ c_2 & r_2 \\ \vdots & \vdots \\ c_N & r_N \end{bmatrix} \quad (1)$$

여기서  $r, c$  및  $N$ 은 각각 주어진 입력 이미지의 행, 열 및 총 픽셀 수를 나타냅니다. PCA 알고리즘에서 평균과 공분산 행렬을 계산하는 것은 데이터 확산의 최대 분산 방향을 나타내는 첫 번째 구성 요소를 찾는 데 필수적입니다. 평균은 다음과 같이 계산할 수 있습니다.

$$m_z = \frac{1}{N} \sum_{i=1}^N Z_k \quad (2)$$

공분산은 일반적인  $2 \times 2$  행렬이며 다음과 같이 계산할 수 있습니다.

$$cz = \frac{1}{N} \sum_{k=1}^N (Z_k - m_z)(Z_k - m_z)^T \quad (3)$$

고유값과 고유벡터는 각각 분산의 크기와 방향을 나타냅니다. 고유값 문제는 다음과 같이 풀 수 있습니다.

$$(cz - \lambda I)e = 0 \quad (4)$$

여기서  $e, \lambda$  및  $I$ 는 각각 고유값, 고유벡터 및 항등행렬을 나타냅니다. 식에서  $e$ 에 대한 해를 찾으려면 (4),  $(cz - \lambda I)$ 는 다음과 같이 비특이 행렬이어야 합니다.

$$\det(cz - \lambda I) = 0 \quad (5)$$

행렬식의 해는 다음과 같은 2차 방정식을 제공합니다.

$$\lambda^2 - \text{tr}(cz)\lambda + \det(cz) = 0 \quad (6)$$

고유값은 2차 방정식을 풀어서 계산할 수 있습니다.

다음과 같은 2도 공식:

$$\lambda_{2,1} = \frac{\text{tr}(cz) \pm \sqrt{\text{tr}(cz)^2 - 4\det(cz)}}{2} \quad (7)$$

$$\text{tr}(cz) = x + y \text{ 및 } |cx| = xy - b \quad (8)$$

공분산 행렬  $cz$ 의 고유 벡터(각각  $2 \times 1$  벡터) 다음과 같이 계산할 수 있는 고유값에 해당합니다.

$$u = \frac{1}{b^2 + (\lambda_1 - x)^2} \begin{bmatrix} b \\ \lambda_1 - x \end{bmatrix} \quad (9)$$

$$v = \frac{1}{b^2 + (\lambda_2 - x)^2} \begin{bmatrix} b \\ \lambda_2 - x \end{bmatrix} \quad (10)$$

$X$ 는 열이  $u = [x_{11}, x_{21}]$  및  $v = [x_{12}, x_{22}]$ 인 행렬입니다. 첫 번째 벡터  $u$  및  $v$ 는 형성하는 가장 큰 고유값에 해당하며 두 번째 벡터는 다음 고유값에 해당합니다.

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \quad (11)$$

2 단계: 행렬  $X$ 를 사용하여 주어진 잘못 정렬된 이미지의 회전 각도는 다음과 같이 계산할 수 있습니다.  $\text{tr}(X)$

$$\theta = \cos^{-1} \left( \frac{\text{tr}(X)}{\sqrt{\text{tr}(X)^2 - 2\det(X)}} \right) \quad (12)$$

여기서  $\text{tr}(X) = x_{11} + x_{22}$ 입니다. 원본 이미지의 표준 방향을 찾기 위해 회전 각도는 다음과 같이 계산됩니다.

$$\text{회전} = \begin{cases} \theta & \text{if } \theta = 0 \text{ 회전 예 그렇지 않으면 아니요} \end{cases} \quad (13)$$

여기서  $\theta$ 는 회전된 객체의 실제 면을 나타냅니다.

3단계:  $\theta$ 는 회전된 이미지를 변환하기 위해 원하는 각도입니다. 표준 방향으로 변환하면 다음과 같이 계산할 수 있습니다.

$$\theta = (x_{11} \times x_{21})\theta \quad (14)$$

PCA는 이미지를 기본 확산 방향으로 공간적으로 정렬했습니다. 잘못 정렬된 이미지는 식에서 계산된 각도로 이미지를 회전하여 진정한 표준 방향으로 다시 변환됩니다. (14) 쌍선형 보간법 사용. 이미지 정렬 결과는 그림 3에 나와 있습니다.

2) 이미지 향상 및 노이즈 제거 대부분의 점자 이미지는 노이즈에 의해 과장되며 노이즈 이미지에 대해 학습된 딥 모델은 인식 성능이 저하됩니다. 따라서 심층 CNN 모델에 이미지를 제공하기 전에 이미지 품질을 향상시켜야 할 필요성이 높습니다.

이미지 향상 단계는 다음과 같습니다.

1 단계: 객체 ROI 추출(점자 이미지의 경우 도트 패턴)의 첫 번째 단계에서 노이즈가 입력된 이미지  $I$ 를 Wiener 필터[58]로 필터링합니다. 노이즈가 입력된 이미지에서 노이즈를 매끄럽게 함과 동시에 블러 효과를 반전시킵니다. 이것

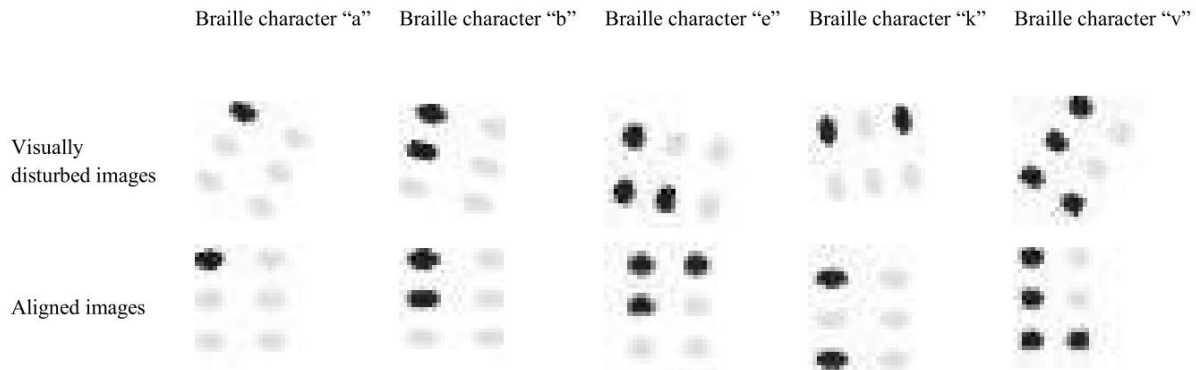


그림 3. 영어 점자 문자 데이터셋에 대한 표준 방향 결과[41]. 시각적으로 교란된 점자 이미지 및 해당 정렬 이미지 이미지는 세로 축을 중심으로 회전합니다.

필터링 프로세스는 도트 영역을 향상시키고 매끄러운 경계를 유지합니다.

2 단계 : 다음 단계에서는 히스테리시스 임계값[59]이 적용됩니다. 임계값 설정 후 얻은 이진 이미지는 다음과 같이 표현할 수 있습니다.

$$I_w = \begin{cases} I(x, y) < T_1 \text{ 인 경우 } 0, \\ I(x, y) > T_2 \text{ 인 경우 } 1 \\ 0, T_1 \leq I(x, y) \leq T_2 \text{ 이고 } I(x, y) \text{가 레이블이 0인 픽셀의 이웃인 경우} \\ 1, T_1 \leq I(x, y) \leq T_2 \text{ 이고 } I(x, y) \text{가 레이블이 1인 픽셀의 이웃인 경우} \end{cases} \quad (15)$$

여기서  $T_1$  및  $T_2$ 는 상한 및 하한 임계값을 나타냅니다.

실험에서 상위 임계값  $T_1$ 을 0.9로 설정하고 하위 임계값  $T_2$ 를 0.45로 설정했습니다.

3 단계 : 이진 이미지는 디스크 모양의 구조 요소를 사용하여 재구성하여 형태학적 개방으로 처리됩니다[60].

$$\phi(I) = I \oplus \text{ps}(I_w) \quad (16)$$

여기서  $\text{ps}(I_w)$ 는 구조 요소  $s$ 와 마스크 이미지  $I_w$ 의 재구성 연산자를 나타냅니다. 반면 마스크 이미지는 다음과 같이 개방 프로세스 후에 얻을 수 있습니다.

$$I_w = (IS) \oplus S \quad (17)$$

4 단계 : 그런 다음 마스크 이미지  $I_w$ 는 형태학적 침식 및 확장 작업을 사용하여 이미지  $I_0$ 를 출력하도록 처리됩니다 [60]. 그레이스케일 침식은 다음과 같이 설명됩니다.

$$(I_w S)(x, y) = \min_{(x+x', y+y') \in V_s} I_w(x+x', y+y') \quad (18)$$

여기서  $r$ 과 구조 요소 사이의 이진 침식을 나타냅니다.  $V_s$ 는 구조 요소의 영역입니다.

5단계: 마찬가지로 그레이스케일 확장은 다음과 같이 작성됩니다.

$$(I_w \oplus S)(x, y) = \max_{(x-x', y-y') \in V_s} I_w(x-x', y-y') \quad (19)$$

위너 필터링, 역치화, 모폴로지 연산의 결과는 그림 4와 같다. 그림 4에서 위너 필터는 이미지가 흐려지는 경우에도 부가적인 노이즈를 제거하여 핵심 영역을 강조하고 이미지 품질을 향상시키는 것을 볼 수 있다. 낮은 강도.

형태학적 작업은 불필요한 개체를 제거하고 매끄러운 경계를 유지하는 데 도움이 됩니다. 또한 침식 및 확장은 최종 ROI 이미지에서 불필요한 개체를 제거하는 데 도움이 됩니다. 이미지 정렬 및 향상의 궁극적인 목표는 이미지 품질을 개선하는 것입니다. 실험 분석을 통해 우리는 이미지 전처리가 인식 효율성 측면에서 심층 CNN 모델의 성능을 크게 향상시킨다는 것을 발견했습니다. 이미지 정렬 및 향상 단계에 대한 자세한 내용은 알고리즘 1에 나와 있습니다.

#### 나. 제안하는 이미지 분류 방법

본 장에서는 제안하는 심층 CNN 모델을 체계적으로 설명하고 다양한 활용 기법에 대해 논의한다. 컨볼루션 신경망은 패턴 인식 작업에 주로 적용되는 인공 신경망(ANN)의 한 형태입니다. 다른 ANN 네트워크와 달리 CNN은 공간 유형의 커널과 보폭을 사용하여 시간 및 공간 세부 정보를 유지하는 반면 ANN은 정보를 1차원 메트릭으로 억제합니다. 많은 연구자들이 이미지 처리 분야에서 CNN을 사용하여 강력한 성능을 얻었습니다. Deep CNN은 주로 대안, 컨볼루션, 최대 풀링 및 저수준 표현에서 고수준 세부 정보를 학습하는 완전 연결 계층으로 구성됩니다[42]. 이러한 계층은 보폭, 커널 크기 및 패딩과 같은 여러 하이퍼 매개변수를 조정하여 훈련됩니다.

본 논문에서는 피쳐 추출과 분류 레이어라는 두 개의 주요 레이어로 구성된 심층 CNN 모델을 설계했습니다. 특징 추출을 위해 이전에 제안된 심층 CNN 네트워크인 VGG16[43], ResNet50[45], InceptionV3[44] 및 DensNet201[46]을 선택했습니다. 각각의 원래 네트워크에서 마지막 몇 개의 모듈은 IRB 모듈[57]로 대체됩니다.

우리는 VGG16[43], ResNet50[45], InceptionV3[44],



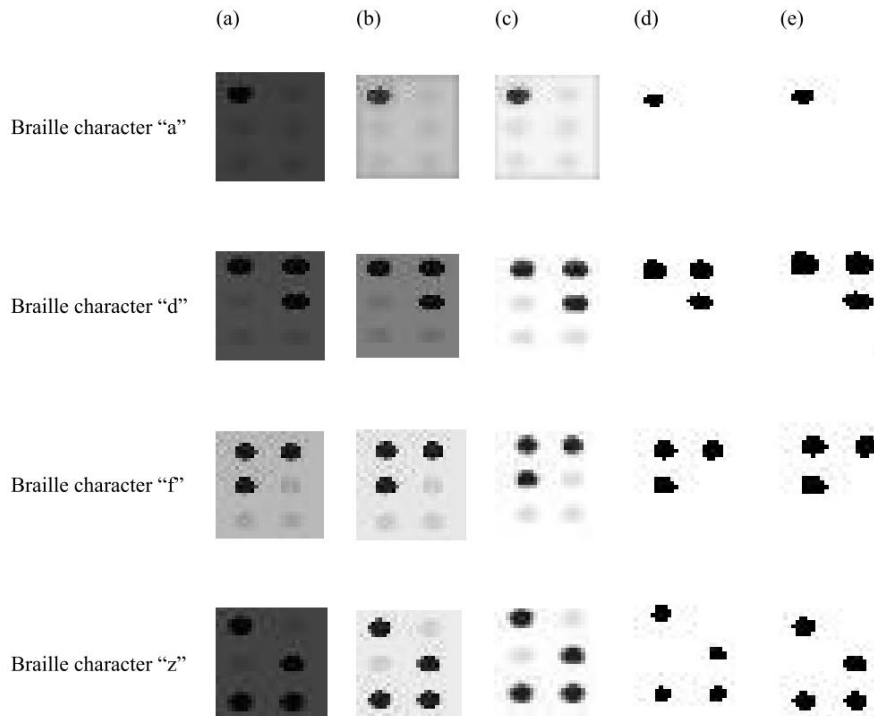


그림 4. (a) 원본 이미지, (b) 위너 필터로 필터링, (c) 히스테리시스 임계값 후, (d) 형태학적 재구성 후, (e) 형태학적 플러드 필, 즉 침식 후 팽창. 이미지는 English Braille 데이터 세트에서 가져온 것입니다.

#### 점자 이미지 전처리 알고리즘 1

입력: 노이즈 이미지 I

출력: 처리된 이미지 I<sub>o</sub>

절차:

이미지 정렬

##입력 이미지 제공

X =

x11 x12

x21 x22

$\theta = \cos^{-1} \frac{\text{tr}(X)}{2}$

$\theta = (x_{11} \times x_{21}) / \theta$

/eigenvalues 및 해당 고유 벡터는 다음에서 계산됩니다.

/오정렬된 이미지의 회전 각도가 계산됩니다. /오정렬된 이미지는  $\theta$ 를 사용

하여 표준 방향으로 다시 회전됩니다.

이미지 향상

Iw = Wiener(I, 잡음)

$I(x, y) = \frac{1}{N} \sum_{i=1}^N I(x_i, y_i)$  경우 2 인

Iw =

0,  $T_1 \leq I(x, y) \leq T_2$  이고  $I(x, y)$ 가 레이블이 0인 픽셀의 이웃인 경우

1,  $T_1 \leq I(x, y) \leq T_2$  이고  $I(x, y)$ 가 레이블이 1인 픽셀의 이웃인 경우

$\phi(Iw) = Iw - \text{ps}(Iw) / \text{재구성에 의한 형태학적 개방}$

$(Iw \otimes S)(x, y) = \text{최소 } Iw(x + x', y + y') \mid (x, y) \in V_s$  /형태적 침식

$(Iw \ominus S)(x, y) = \text{최대 } Iw(x - x', y - y') \mid (x, y) \in V_s$  /형태학적 팽창

/히스테리시스 임계값

DenseNet201 [46]. 백본으로 VGG16의 경우 마지막 3개의 컨볼루션 레이어가 IRB 모듈로 대체됩니다. ResNet50의 경우 마지막 6개의 컨볼루션 레이어가 IRB 모듈로 대체됩니다. 백본으로 InceptionV3의 경우, 마지막

인셉션 모듈은 IRB 모듈로 대체됩니다. 기본적으로 DenseNet201의 주요 구조는 4개의 서로 다른 고밀도 블록과 전환 레이어로 구성됩니다. 전환 레이어( $1 \times 1$  컨볼루션 다음에  $2 \times 2$  평균 풀링)는 다음과 같습니다.

두 개의 인접한 고밀도 블록 사이에 사용하여 연결합니다.

각 고밀도 블록은 각각 6, 12, 48, 32개의 병목 현상 계층으로 구성됩니다. 각 병목 현상 계층에서 BN-ReLU-1×1 Conv는 BN-ReLU-3×3 Conv 전에 사용되며 여기서 BN은 Batch Normalization을 나타냅니다. DenseNet201의 경우 IRB 모듈을 사용하여 마지막 dense 블록에서 10개의 Bottleneck 레이어를 제거했습니다.

설계된 심층 CNN 모델(DenseNet201 백본 포함)의 네트워크 아키텍처는 그림 5에 나와 있습니다. IRB 모듈에서 깊이 방향 컨볼루션 계층 전후에 1×1 포인트 방식 컨볼루션이 추가되었습니다. 깊이별 컨볼루션에서는 출력 특징 맵 (Cout)의 채널 수가 입력 특징 맵 (Cin)과 동일하기 때문입니다. 점별 컨볼루션 레이어는 출력 피쳐 맵의 차원을 증가시킵니다. 특징 맵이 n배 확장됨에 따라 출력 채널 수가 n×Cin으로 증가합니다. 이러한 방식으로 네트워크 작업은 고차원 공간에서 더 많은 차별적 특징을 추출하고 네트워크 성능을 향상시킵니다. 두 번째 pointwise convolution 레이어를 사용하면 출력 채널이 입력 채널과 동일하고 특징 정보가 채널에 집중되도록 출력 차원이 줄어듭니다.

반전 잔차 블록의 아키텍처는 그림 6에 나와 있습니다.

또한 수정된 네트워크는 컨볼루션, 배치 정규화, ReLU 활성화 함수와 전역 평균 풀링 레이어로 구성됩니다. IRB 모듈에서는 512개의 필터를 사용했으며 비교 분석을 위해 하이퍼파라미터 확장 계수 n을 1, 2, 3으로 설정했습니다. 제안된 모델에서 제안된 수정은 계산 비용을 줄이고 더 적은 네트워크 매개변수로 이어집니다.

분류 과정은 주어진 각 클래스의 확률 값을 계산하는 데 사용되는 Softmax 함수를 중심으로 수행했습니다. Softmax를 사용한 분류는 다음과 같이 공식화됩니다.

$$\sigma = \frac{e^{y_j}}{\sum_{j=1}^N e^{y_j}} \quad (20)$$

여기서  $y_i$ 는 softmax에 대한 입력,  $j$ 는 총 범주 수,  $\sigma$ 는 정규화된 Softmax 출력 확률입니다. 이미지에 예측된 Softmax 클래스 확률 값이 가장 높은 클래스 레이블이 할당됩니다. Softmax 클래스 확률은 다음과 같이 계산됩니다.

$$o = \text{최대 인수}(\sigma) \quad (21)$$

인식 모델에서 Softmax의 범주형 교차 엔트로피 손실 함수는 실제 출력과 원하는 출력 사이의 근접성을 찾는 데 사용됩니다. 손실 함수는 다음과 같이 계산됩니다.

$$J = \frac{1}{N} \sum_{t=1}^N -\log \frac{e^{y_{oj}}}{\sum_{j=1}^N e^{y_j}} \quad (22)$$

여기서  $N$ 은 훈련 샘플의 총 수입니다. 네트워크 훈련은 손실을 최소화하기 위해 파라미터 공간에서 최적의 솔루션 그룹을 찾는 최적화 프로세스로 간주됩니다. 손실 함수의 기울기가 계산되고 훈련 손실이 가장 낮은 값으로 최소화될 때까지 가중치 매개변수가 업데이트됩니다.

1) 반전을 통한 테스트 시간 최소화  
잔여 블록

커널 크기가  $k \times k$ 인 컨볼루션 레이어가 입력을 받도록 합니다.

이전 레이어에서 Cin을 생성하고 크기 Cout의 출력 피쳐 맵을 생성합니다. 표준 컨볼루션 레이어[61]에 포함된 시간 계산은 다음과 같이 계산할 수 있습니다.

$$\text{Ostd Cin} \cdot k^2 \cdot \text{Mout} \cdot \text{Nout} \cdot \text{Cout} \quad (23)$$

여기서 Mout, Nout 및 Cout은 출력 기능 맵의 공간 차원입니다. 예를 들어 첫 번째 컨볼루션 레이어의 시간 계산은 다음과 같이 작성됩니다.

$$\text{Ostd} (C0.k^2 \cdot M1.N1.C1) \quad (24)$$

방정식 24는 CNN 아키텍처의 시간 복잡도가 주로 입력 이미지의 공간 차원, 커널 수 및 사용된 커널 크기에 의존한다는 것을 보여줍니다. 우리가 제안한 아키텍처에서 IRB와 작은 크기의 이미지를 사용하는 것이 계산 복잡도를 줄이는 데 크게 기여합니다.

예를 들어, 입력 이미지의 크기가  $28 \times 28 \times 3$ 에서  $14 \times 14 \times 3$ 으로 줄어들면 첫 번째 컨볼루션 레이어의 시간 복잡도는 75%로 줄어들고 다음과 같이 계산됩니다 (식 24 사용).

$$\begin{aligned} \text{Ostd} (64 \times 3^2 \times 28 \times 28 \times 3) \\ \rightarrow \text{Ostd} (64 \times 3^2 \times 14 \times 14 \times 3) \quad (25) \end{aligned}$$

또한 IRB[57]를 사용하면 제안된 모델의 계산 비용을 효과적으로 줄일 수 있습니다. IRB는 전통적인 컨볼루션을 깊이별 컨볼루션과  $1 \times 1$  포인트별 컨볼루션으로 분리하여 계산 비용을 줄이는 깊이별 분리 가능 컨볼루션입니다. depthwise separable convolution의 계산 비용은 deep convolution의 계산량과  $1 \times 1$  convolution의 합으로 정의된다.

Depthwise separable convolution의 계산 비용은 다음과 같이 정의됩니다.

$$\text{Odpt Cin} \cdot f^2 \cdot \text{Mout} \cdot \text{Nout} + \text{Cin} \cdot \text{Cout} \cdot \text{Mout} \cdot \text{Nout} \quad (26)$$

표준 컨볼루션에 대한 깊이별 분리 가능 컨볼루션의 계산 비용 비율은 다음과 같이 계산할 수 있습니다.

$$\begin{aligned} \text{cr} = \text{Odpt Cin} \cdot f^2 \cdot \text{Mout} \cdot \text{Nout} + \text{Cin} \cdot \text{Cout} \cdot \text{Mout} \cdot \text{Nout} \\ \times \text{Ostd Cin} \cdot \text{에프}^2 \cdot \text{Mout} \cdot \text{Nout} \cdot \text{Cout} \quad (27) \end{aligned}$$

$$\text{cr} = \frac{1}{\text{쿼드트}} + \frac{1}{\text{에프}} \cdot \frac{2}{\text{에프}} \quad (28)$$

여기서 방정식 (26)과 (28)은 곱셈 연산을 곱셈과 덧셈 연산으로 바꾼 깊이 분리 컨볼루션의 장점을 나타낸다. 예를 들어, 첫 번째 레이어에서  $14 \times 14$  픽셀의 다운샘플링된 입력 점자 이미지를 사용하면 표준 컨볼루션에서  $64 \times 3 \times 3 \times 3$  커널이  $12 \times 12$  번 이동합니다(즉,  $64 \times 3 \times 3 \times 3 \times 12 \times 12 = 248,832$ ). 곱셈). 분리 가능한 컨볼루션과 달리 깊이 방향 컨볼루션에서는  $3 \times 5 \times 5 \times 1$  커널이  $12 \times 12$  번 이동하고(즉,  $3 \times 3 \times 3 \times 12 \times 12 = 3,888$  곱셈) 점 방향 컨볼루션에서는  $64 \times 1 \times 1 \times 3$  커널 이동

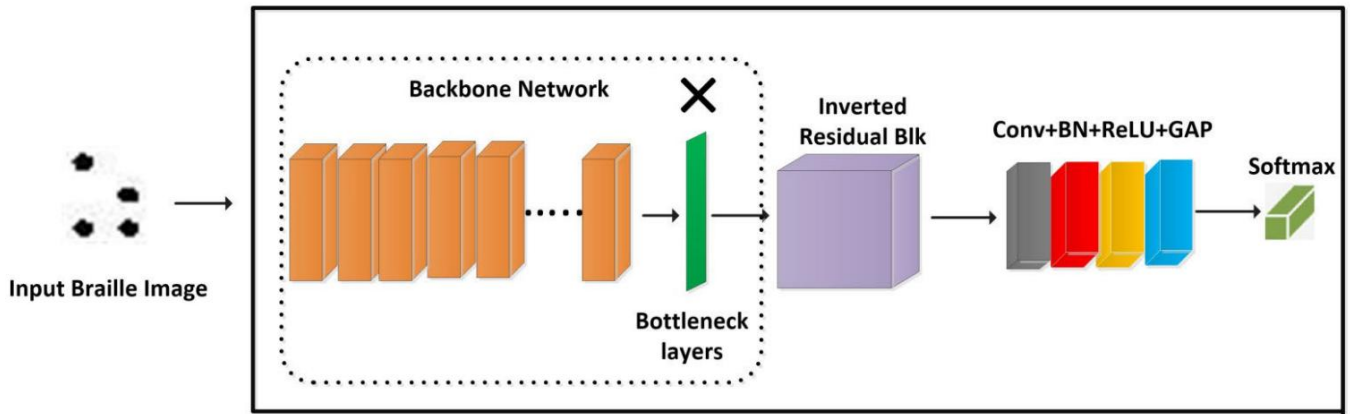


그림 5. 제안된 심층 CNN 모델(DenseNet201 백본 포함)의 네트워크 아키텍처. 백본 네트워크에서 DenseNet의 마지막 고밀도 블록에 있는 10개의 병목 레이어가 반전된 잔차 블록으로 대체됩니다.

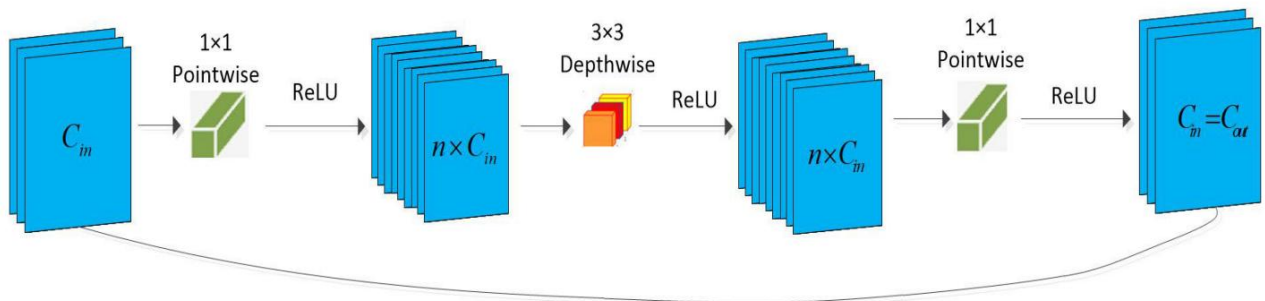


그림 6. 반전된 잔차 블록.

$12 \times 12$ 배(즉,  $64 \times 1 \times 1 \times 3 \times 12 \times 12 = 7,648$  곱셈). 깊이별 컨볼루션과 포인트별 컨볼루션의 곱셈을 합하면 31,536 곱셈이 됩니다. 이러한 계산은 깊이 분리 컨볼루션을 사용하면 첫 번째 컨볼루션 레이어의 시간 복잡도가 87.3% 감소함을 보여줍니다.

**III. 실험 설정 및 결과 분석** 이 섹션에서는 수정된 모델의 아키텍처, 실험 설정, 사용된 데이터 세트 및 출력 결과에 대해 설명합니다. 이미지 전처리 알고리즘은 MATLAB R2018a 프로그램으로 구현됩니다. 반면, CNN 알고리즘은 12GB 메모리의 NVIDIA Tesla M40 GPU 1개가 있는 2.4GHz Intel(R) Xeon(R) E5-2630 CPU에서 TensorFlow[62] 및 Keras[63] 라이브러리로 구현됩니다. 첫째, 사용된 다양한 전처리 기술이 성능에 미치는 영향을 조사합니다. 주요 목표는 이전에 제안된 다양한 백본 네트워크가 출력 인식 성능에 미치는 영향을 확인하는 것입니다.

#### 가. 통계적 측정

제안된 모델의 인식 성능을 평가하기 위해 다음과 같은 조치를 취합니다.

$$\text{감도} = \frac{TP}{TP + FN} \quad (29)$$

$$\text{퍼세션} = \frac{TP}{TP + FP} \quad (30)$$

$$F \text{ 점수} = 2 \times \frac{\text{percesion} \times \text{민감도}}{\text{percesion} + \text{민감도}} \quad (31)$$

$$\text{예측 정확도} = \frac{TP + TN}{TP + TN + FP + FN} \quad (32)$$

여기서 TP, FP, TN&FN 은 각각 참양성, 거짓양성, 참음성, 거짓음성을 나타냅니다.

멸망의 엔트로피[64]:

$$H = - \sum p(y_i, D) \log p(y_i, D) \quad y_i \in Y \quad (33)$$

픽크 신호 대 잡음비(PSNR)[65]는 다음과 같습니다.

$$\text{PSNR} = 20 \log \text{MAXI} \sqrt{\text{MSE}} \quad (34)$$

여기서 MAXI 는 이미지의 최대 강도 값을 나타내고 MSE 는 원본 이미지와 노이즈 오염 이미지 간의 평균 제곱 오차입니다.

#### 나. 데이터셋

본 논문에서는 제안하는 방법의 점자 문자 검출 성능을 공개적으로 이용 가능한 두 개의 점자 데이터셋, 즉 영어 점자 문자[41]와 양면 점자 이미지 데이터셋[40]으로 확인한다.

데이터 세트에 대한 자세한 설명은 후속 단락에서 제공됩니다.



**표 1.** DenseNet201을 백본으로 사용하여 제안된 방법에 대한 이미지 전처리(a: 정렬, e: 향상), 데이터 증대(dim: 밝기 보정, rot: 회전, whs: 폭 높이 이동) 및 다운샘플링의 성능에 미치는 영향 분석. 결과는 영어 점자 데이터 세트에서 평가됩니다.

Experiments	Techniques	Precision (%)	Sensitivity (%)	F score (%)	Entropy (%)
1	Proposed method with original images	88.7	87.9	88.2	32.3
2	Proposed method+A	89.7	91.5	90.5	20.5
3	Proposed method+E	92.4	91.9	92.1	20.1
4	Proposed method+dim+ rot+ whs	93.9	93.7	92.2	19.7
5	Proposed method+downsampling	91.9	91.3	92.2	19.8
6	<b>Proposed method+(rot+whs+dim)+E+A+downsampling</b>	<b>95.8</b>	<b>94.7</b>	<b>95.2</b>	<b>19.2</b>

**표 2.** 영어 점자 데이터 세트를 사용한 실험의 하이퍼 매개변수 설정.

Experiments	Learning rate	Batch size	Epochs	Trining Time
1	0.0001	256	60	1950s
2	0.0003	64	90	2890s
3	<b>0.0002</b>	<b>8</b>	<b>150</b>	<b>4800s</b>

#### 1) 영어 점자 데이터셋 [41]

이 점자 이미지 데이터 세트는 오픈 소스 Kaggle 저장소[41]에서 공개적으로 사용할 수 있습니다. 26자를 포함한 1560개의 이미지로 구성되어 있습니다. 이러한 점자 문자의 이미지는  $28 \times 28 \times 3$  픽셀의 해상도를 가진 흑백 컬러 스케일을 갖는 그레이스케일 3채널 이미지입니다. 데이터 세트를 교육, 검증 및 테스트 세트로 분할합니다. 점자 영어 문자 데이터셋의 약 20%와 10%는 테스트 세트와 검증 세트에서 각각 선택됩니다. 나머지 이미지는 모델 교육에 사용됩니다. 훈련 데이터셋의 크기는 상대적으로 작기 때문에 회전, 너비 높이 이동 및 밝기 보정을 통해 원본 이미지의 증강 버전이 생성됩니다. 이러한 데이터 증가는 데이터 크기를 60배, 즉 3 증가  $\times$  20 다른 증가 값으로 늘렸습니다.

#### 2) 중국어 양면 점자 이미지(DSBI) 데이터 세트 [40]

이 점자 이미지 데이터 세트는 오픈 소스 Github 리포지토리[40]에서 공개적으로 사용할 수 있습니다. 그것은 여러 점자 책에서 얻은 114개의 양면 점자 이미지를 포함합니다. 이 데이터 세트의 증강 버전도 회전, 너비 높이 이동 및 원본 이미지의 밝기 보정으로 생성됩니다. 이 데이터 세트에서 각 점자 이미지는 해상도가 다릅니다. 이 데이터 세트를 교육, 검증 및 테스트 세트로 나누었습니다. 각 세트는 각각 70, 14 및 30개의 이미지로 훈련, 검증 및 테스트 부를 구성하기 위해 비례적으로 샘플링됩니다. 우리는 직장 점자 문자와 역 점자 문자 모두에 대한 결과를 평가했습니다. 이 데이터 세트는 심층 CNN 모델의 전체 교육을 제공할 만큼 충분히 크지는 않지만 문제에 대한 다양한 방법을 비교할 수 있습니다.

#### 다. 영어 점자 데이터셋에 대한 평가

이 섹션에서는 영어 점자 데이터 세트 [41]에서 얻은 결과를 보여줍니다.

##### 1) ABLATION EXPERIMENTS 전처

리, 데이터 증가 및 이미지 다운 샘플링 기술의 성능에 미치는 영향을 확인하기 위해 개선된 모델(Backbone: DenseNet201)을 사용하여 여러 실험을 수행했으며 결과는 표 1에 나와 있습니다.

첫 번째 실험에서 모델은 원본 이미지로 훈련됩니다(전처리 없음). 두 번째 실험에서 모델은 PCA 알고리즘을 사용하여 정렬된 이미지로 학습됩니다('A'는 이미지 정렬을 나타냄). 세 번째 실험에서 모델은 이미지 향상 기술로 전처리된 이미지로 학습됩니다('E'는 이미지 향상을 나타냄). 유사하게, 네 번째 실험에서 학습 프로세스는 데이터 확대(rot: 회전, whs: 너비 높이 이동, dim: 밝기 보정) 기술을 통합하여 수행됩니다. 또한 다섯 번째 실험에서는 이미지 다운샘플링이 모델 성능에 미치는 영향을 분석합니다. 입력 이미지의 해상도는 쌍선형 보간법을 사용하여  $28 \times 28 \times 3$  픽셀에서  $14 \times 14 \times 3$  픽셀로 다운샘플링됩니다.

여섯 번째 실험에서는 이미지 전처리, 데이터 증대 및 다운 샘플링 기술을 모두 포함하여 학습 과정을 반복합니다.

표 1에서 볼 수 있듯이 이미지 전처리는 이미지 품질을 향상시키고 인식 효율성 측면에서 심층 CNN 모델의 성능을 크게 향상시킵니다. 특히, 이미지 정렬 및 향상 단계를 통해 점자 인식에서 높은 정확도를 얻을 수 있습니다. 실험 분석에서 우리는 데이터 증대의 통합이 분류 정확도의 일관된 개선을 나타냄을 발견했습니다. 이미지 다운샘플링이 인식 성능에 미치는 영향도 관찰할 수 있습니다. 작은 크기의 입력 이미지는 인식 정확도에 큰 영향을 미치지 않음을 알 수 있습니다. 그러나 다운샘플링이 있는 매개변수는 다운샘플링이 없는 매개변수보다 낮습니다. 전체 결과는 모든 이미지 전처리, 증강 및 다운샘플링 기술이 다음일 때 모델이 우수한 성능을 달성했음을 보여줍니다.

표 3. 수정된 네트워크(다른 백본 사용)의 f 점수 및 매개변수와 영어 점자 데이터 세트의 해당 원본 네트워크 비교. F 점수와 매개변수는 확장 계수 n의 다른 값을 사용하여 비교됩니다.

Methods	Entropy (%)	n=1		n=2		n=3	
		F score (%)	Parameters	F score (%)	Parameters	F score (%)	Parameters
<b>Proposed method (Backbone:DenseNet201)</b>	<b>19.2</b>	<b>95.2</b>	<b>18,430,747</b>	<b>95.4</b>	<b>26,706,764</b>	<b>95.9</b>	<b>32,080,252</b>
Proposed method (Backbone:InceptionV3)	20.6	94.3	21,611,487	94.8	28,830,234	95.1	33,902,231
Proposed method (Backbone:ResNet50)	21.4	89.8	29,430,123	90.1	34,50,231	91.6	39,630,456
Proposed method (Backbone:VGG16)	21.3	87.3	141,231,342	87.7	149,361,165	88.3	156,410,123

존경받는. 결과적으로 모델은 점자 인식에 대해 전체 F 점수 95.2%를 획득했습니다.

또한 서로 다른 하이퍼 매개변수 설정으로 세 가지 실험을 수행합니다. 하이퍼 매개변수는 심층 CNN 모델의 적절한 교육에 도움이 됩니다. 각 실험의 하이퍼파라미터 설정과 훈련시간은 Table 2와 같다. 우리는 반복적으로 모델을 수정하고 모델의 하이퍼파라미터를 조정하여 더 나은 성능을 얻었다. rmsprop(Root Mean Square Propagation) 옵티마이저를 사용하여 0.0002의 학습률과 8의 배치 크기로 150 epoch 동안 분류기를 학습한 세 번째 실험에서 최상의 성능을 얻었습니다. 이러한 하이퍼 매개변수 값은 학습 데이터에 대한 여러 실험을 거쳐 휴리스틱 방식으로 선택됩니다.

## 2) 다른 백본을 사용한 실험 설계된 모델의 성능을 추가로 평

가하기 위해 확장 계수 n을 변경하여 다른 백본 네트워크를 사용하여 결과를 평가합니다. 백본 네트워크가 다른 우리가 제안한 모델의 F 점수와 매개변수는 표 3에 보고되어 있습니다.

결과는 다른 백본을 가진 우리의 방법이 다른 정도의 개선을 얻는다는 것을 보여줍니다. 일반적으로 DenseNet201과 같은 백본을 사용하는 방법은 대부분의 경우 가장 높은 F 점수를 얻습니다. 1, 2 또는 3의 다른 확장 요인을 사용하는 것은 F 점수에 거의 영향을 미치지 않으며 모든 n 값에 대해 거의 유사하다는 점에 유의하는 것이 중요합니다. 그럼에도 불구하고 계산 매개변수는 n = 1일 때보다 n = 2 또는 n = 3일 때 훨씬 높습니다. 깊은 컨볼루션 계층에서 n 값이 증가하면 매개변수가 급격히 증가합니다. 따라서 실험에서 n = 1을 선택했습니다. 이는 매개변수를 적게 할 뿐만 아니라 상당한 분류 정확도를 제공합니다.

테스트 데이터 세트에서 서로 다른 백본 네트워크 [43]-[46]로 평가된 성능을 비교하기 위해 제안된 모델의 정밀도-재현율 곡선은 그림 7에 나와 있습니다. 서로 다른 백본과 성능을 비교한 후(n = 1 사용), 수정된 모델이 DenseNet201로 최상의 분류 결과를 달성했음을 발견했습니다. 평가된

결과는 우리의 방법이 입력 점자 이미지 품질을 향상시키고 모델 매개변수를 효과적으로 감소시킨다는 것을 보여줍니다. 또한 점자 문자의 강력한 인식 정확도를 달성할 수 있는 이점을 제공합니다.

3) 최신 기술과의 분류 정확도(%) 비교 이 섹션에서는 우리의 결과를 이미지 분류에 사용되는 최신 기술과 정교하게 연관시킵니다. 백본

DenseNet201[46](확장 계수 n = 1 사용)을 사용한 방법의 성능은 원래 InceptionV3[44], ResNet50[45], VGG16[43] 네트워크 및 가장 진보된 DenseNet201 네트워크[46]와도 비교됩니다. 수정된 모델은 원래 네트워크 [43]-[46](표 4에 보고됨)보다 매개변수가 적다는 것을 알 수 있습니다.

원래 네트워크의 일부 모듈을 매개변수를 줄이는 IRB로 교체했기 때문입니다. 성능 분석에서 수정된 방법의 F 점수가 상응하는 원래 InceptionV3, DenseNet201, ResNet50 및 VGG16 모델(표 4에 보고됨)과 비교할 때 상대적으로 더 우수하다는 것도 분명합니다. 전반적으로, 실험 성능은 제안된 접근 방식이 문자 인식 정확도를 향상시킬 뿐만 아니라 파라메트릭 부하를 효과적으로 감소시킨다는 것을 나타냅니다.

엔트로피 측정은 예측의 불확실성을 계산하는 데 사용됩니다. 이는 주로 매우 불확실한 예측에 대해 오분류가 발생했다는 사실을 포착합니다. 따라서 모델의 Aleatoric 불확실성[66]을 측정하기 위해 예측의 엔트로피를 계산했습니다. 최신 모델(표 4 참조)과 비교하여 제안하는 방법(Backbone: DenseNet201)은 예측 엔트로피가 19.2%로 낮습니다.

## D. 시간 분석

딥 러닝 모델 훈련은 가장 어렵고 오랜 시간이 걸립니다. 그러나 실시간 시나리오에서는 자동 분류 모델의 인식 속도를 빠르게 하는 것이 최우선이다 [67,68]. 결과는 우리가 제안한 방법(Backbone:DenseNet201)이 더 적다는 것을 보여줍니다.

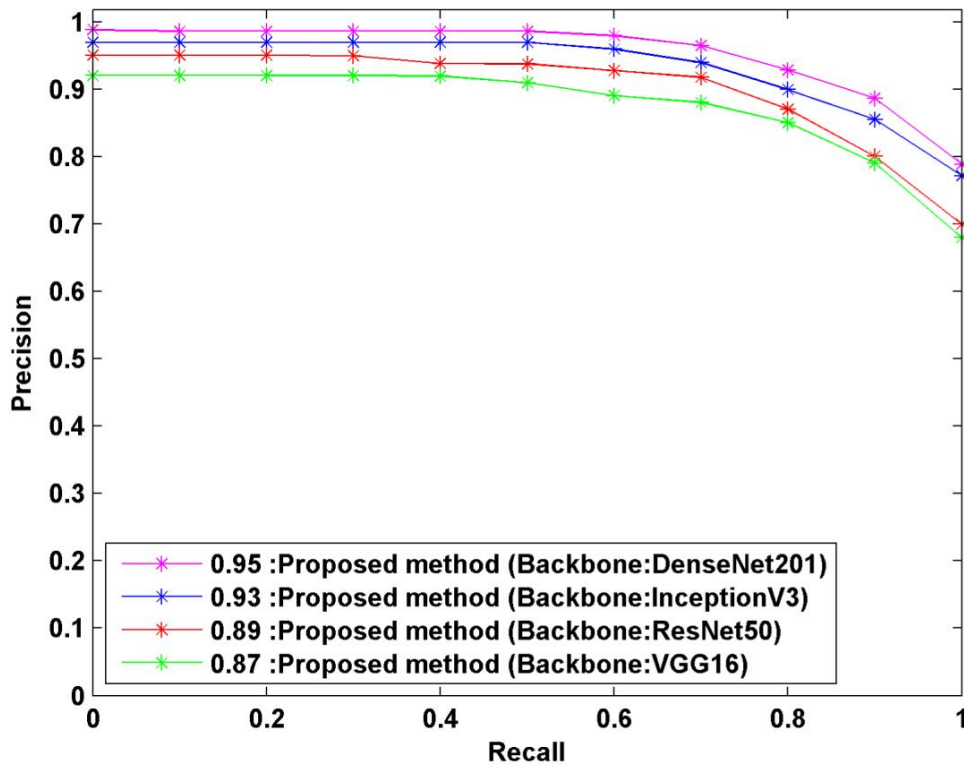


그림 7. 영어 점차에서 서로 다른 백본 네트워크를 사용한 제안 방법의 정밀도-재현율 곡선 테스트 데이터.

표 4. 영어 점차 데이터 세트에 대한 최첨단 모델과의 성능 비교.

Methods	F score (%)	Parameters	Entropy (%)
<b>Proposed method (Backbone:DenseNet201)</b>	<b>95.2</b>	<b>18,430,747</b>	<b>19.2</b>
DenseNet201	92.9	18,610,731	23.5
InceptionV3	92.7	21,852,176	26.4
ResNet50	87.5	29,623,164	33.1
VGG16	85.9	198,231,342	32.7

원본 DenseNet201 모델과 비교한 매개변수입니다.

이는 원래 DenseNet201 모델에서 원래 병목 현상 계층에 비해 매개 변수가 적은 IRB 모듈( $n = 1$  사용)로 대체되는 모듈이 거의 없기 때문에 계산 비용이 낮아지기 때문입니다. 개선된 버전의 DenseNet, InceptionV3, ResNet50 및 VGG16 네트워크에서 IRB를 사용하면 메모리 요구 사항과 추론 시간이 효과적으로 줄었습니다. IRB에서 확장 계수  $n$  값을 적절하게 사용하면 파라메트릭 로드를 제어하는 메커니즘이 제공됩니다. 또한 이미지 다운샘플링 개념은 계산 비용을 제어하는 데에도 도움이 됩니다. 최적화된 매개변수 세트를 얻을 때까지 학습 프로세스를 여러 번 반복했습니다. 훈련 시간

표 5. 영어 및 DSBi 점차 데이터 세트에 대한 학습 및 테스트 시간.

Methods	English Dataset		DSBi Dataset	
	Train Time (s)	Test Time /Image (s)	Train Time (s)	Test Time /Image (s)
<b>Proposed method (Backbone:DenseNet201)</b>	<b>4800</b>	<b>0.01</b>	<b>5300</b>	<b>0.03</b>
Proposed method (Backbone:InceptionV3)	5700	0.08	6300	0.06
Proposed method (Backbone:ResNet50)	7300	0.15	7900	0.10
Proposed method (Backbone:VGG16)	6200	0.30	6800	0.07

다른 실험에 대해 제안된 모델은 표 5에 나와 있습니다.

또한 개선된 DenseNet, InceptionV3, ResNet50 및 VGG16 네트워크 버전의 전체 훈련 및 테스트 시간은 표 5에 나와 있습니다. 훈련 시간은 모델 훈련 시간이고 테스트 시간은 단일 이미지에 대해 계산된 모델의 추론 시간입니다.

#### E. 입력 섭동에 대한 견고성

이 분석에서는 제안된 모델(백본 DenseNet20)의 기능이 입력 이미지 노이즈 섭동에 대해 확인됩니다. 입력 이미지에 대한 모델의 견고성을 평가하기 위해

표 6. 영어 점자 데이터 세트에 대한 예측 정확도(%) 측면의 견고성 분석. 여기서 m은 서로 다른 소음 수준을 나타냅니다.

Techniques	Original images	Preprocessed images	Preprocessed images contaminated with different noise levels				
			m=1	m=3	m=5	m=7	m=9
DenseNet201	82.1	92.9	92.6	79.1	53.5	39.4	23.6
<b>Our proposed model (DenseNet201)</b>	<b>88.2</b>	<b>95.2</b>	<b>94.3</b>	<b>87.6</b>	<b>72.3</b>	<b>69.7</b>	<b>57.2</b>

섭동, 우리는 이미지를 테스트하고 오염된 노이즈 이미지를 분류기에 입력하기 위해 다양한 수준의 곱셈 가우시안 노이즈를 추가했습니다. 이미지 섭동은 다음과 같이 설명됩니다.

$$\bar{i} = i + m.\text{randn}(c1, c2, 3)$$

(35)

여기서  $i$ 는  $c1 \times c2 \times 3$  크기의 입력 이미지를 나타내고,  $i$ 는 오염된 데이터를 나타내고,  $\text{randn}()$ 은 가우시안 노이즈 함수를 나타냅니다. 다양한 노이즈 레벨  $m = 1, 3, 5, 7, 9, 11$ 을 선택했습니다. 이미지 섭동에서 픽셀 값이  $[0, 255]$ 를 초과할 수 있으므로 초과된 값을 잘라서  $[0, 255]$  범위로 제한했습니다. (i) 원본 입력 이미지, (ii) 전처리된 입력 이미지, (iii) 서로 다른 노이즈 수준으로 오염된 전처리된 이미지를 사용하여 별도의 실험을 수행했습니다. 견고성 테스트 결과에서 표 6과 같이 수정된 네트워크는 모든 유형의 입력 이미지에 대해 원본 DenseNet201 네트워크에 비해 강력한 견고성을 가지고 있다고 결론을 내릴 수 있습니다. 이 분석은 또한 불가피한 노이즈 섭동이 있는 더 넓은 범위의 점자 데이터를 수용하도록 수정된 모델의 강력한 일반화 능력을 보여줍니다.

또한 노이즈 레벨 증가가 PSNR[65] 값에 미치는 영향을 연구했습니다. 전처리된 이미지와 노이즈가 많은 이미지(서로 다른 노이즈 레벨로 오염된 이미지)에 대해 계산된(식 34 사용) PSNR은 표 7에 나와 있습니다. 손상된 노이즈가 이미지 표현의 충실도에 영향을 미치는 것으로 확인되었습니다.

F. 중국어 양면 점자(DSBI) 데이터셋에 대한 평가 설계된 모델의 성능을 더 확인하기 위해 양면 점자 이미지 데이터셋에 대한 성능을 평가했습니다[40].

English Brail 데이터 세트[41]와 유사하게 이미지 전처리 기술로 이미지 정렬 및 향상을 수행했습니다. 영어 점자 이미지의 경우와 동일한 교육 프로토콜을 채택했습니다. 이 데이터 세트에서 성능은 정밀도, 민감도 및 F 점수 측면에서도 계산됩니다. DSBI 데이터 세트에 대한 딥 러닝 모델의 적용은 이전에 발표된 몇 가지 작업[37, 28, 40, 36]에서 보고되었습니다. 알고리즘의 효율성을 이러한 접근 방식과 비교했습니다. 비교 방법은 U-Net[40]과 BraUNet[36] 모델을 접근 방식으로 사용했습니다. U-Net[40]과 BraUNet[36]은 심층 CNN 기반 모델입니다. BraUNet은 수정된 U-Net 아키텍처를 기반으로 보조 포그라운드 분할 작업을 통해 점유 영역을 결정하는 분할 모델입니다.

표 7. 다양한 노이즈 레벨로 오염된 이미지의 PSNR

Datasets	Preprocessed images	Preprocessed images contaminated with different noise levels				
		m=1	m=3	m=5	m=7	m=9
English Braille	34.1	32.0	29.2	26.1	23.4	20.2
DSBI	33.2	31.3	28.2	26.3	23.1	20.2

문자. 그리고 그 결과를 전통적인 기계 학습 계단식 분류기를 사용한 최근의 TS-OB 방법[38]과 비교한다. 결과는 또한 최근에 제안된 점자 인식 방법과 비교된다[28]. [40], [36]의 저자는 도트 감지에서 알고리즘의 정확도를 비교한 반면 [27], [38]에서는 정확도를 문자 및 도트 수준에서 비교했습니다. 공정한 성능 비교를 위해 메서드는 동일한 하드웨어에서 테스트됩니다. 새로 분할된 DSBI 점자 이미지 세트(즉, 70개의 교육, 검증 및 테스트 부품)에서 U-Net[40], BraUNet[36], TS-OB[38] 및 RetinaNet[28] 방법을 재교육하고 테스트합니다. , 14 및 30 이미지). 테스트된 방법의 성능 비교는 표 8에 나와 있습니다. 결과는 우리 모델이 DSBI 데이터 세트에서 문자 수준 감지에서 0.983%의 정확도를 달성했음을 보여줍니다. 우리의 방법은 12GB 메모리의 NVIDIA Tesla M40 GPU 1개가 장착된 2.4GHz Intel(R) Xeon(R) E5-2630 CPU에서 이미지당 약 0.03초의 시간이 걸렸습니다. 우리의 방법은 F 점수 측면에서 놀라운 성능을 얻었으며 인식 속도 측면에서 다른 방법을 능가합니다. 이 점자 문자 인식 속도는 실시간 애플리케이션에서 허용될 수 있습니다.

비교된 모든 방법[40], [38], [28], [36]은 계산 부담을 증가시키는 몇 가지 후처리 전략을 적용했습니다. 우리가 제안한 방식에서 IRB 모듈을 사용하면 파라메트릭 부하를 제어하고 CNN 모델의 메모리 요구 사항과 추론 시간을 효과적으로 줄일 수 있는 메커니즘을 제공합니다.

G. McNEMAR의 통계 테스트

우리의 방법과 이전에 제안된 다른 방법으로 얻은 결과 간의 차이의 통계적 유의성을 평가하기 위해 McNemar의 통계 테스트를 수행했습니다. 이 통계 테스트는 주어진 문제에서 하나의 기계 학습 분류기가 다른 것보다 우수한지 여부를 확인하는 데 널리 사용됩니다. McNemar 검정의 귀무 가설은 두 방법이 동일한 값에 동의하지 않는다는 것을 정의합니다. 그러나 두 방법이 다른 값과 일치하지 않으면 귀무가설이 기각됩니다. 이 테스트에서는 유의 수준을 설정합니다.

표 8. DSBi 데이터 세트에서 최첨단 모델과의 성능 비교.

Techniques	Dot level detection			Character level detection			Test time/image (s)
	Precision	Sensitivity	F score	Precision	Sensitivity	F score	
U-Net [40]	0.9172	0.9811	0.948				
Haar [40]	0.9765	0.9638	0.970				
HOG,SVM [38]	0.9314	0.9869	0.958				15.02
TS-OBR [38]	0.9965	0.9997	0.996	0.9928	0.9996	0.9962	1.45
BraUNet [36]				0.9943	0.9988	0.9966	0.25
RetinaNet [28]	0.9995	0.9986	0.9991	0.9985	0.9978	0.9981	0.18
Our method				0.9841	0.9812	0.9831	0.03

표 9. DSBi 데이터 세트에 대한 통계 테스트.

Techniques	p-value
U-Net [40]	0.0001
Haar [40]	0.0032
HOG,SVM [38]	0.0067
TS-OBR [38]	0.012
BraUNet [36]	0.018
RetinaNet [28]	0.035

알파 = 5%에서. 귀무가설의 수용은 p-값이 알파값보다 크다고 가정합니다. 실험 분석 중에 얻은 p-값은 표 9에 나와 있습니다.

얻은 p-값은 5%의 알파 값보다 큼니다. 주어진 테스트에 대한 귀무 가설은 불일치에 차이가 없음을 증명합니다. 테스트된 방법은 오류 비율이 비슷합니다. 귀무 가설은 분류 방법이 유사한 성능 메트릭 값을 가지며 모든 차이는 우연히 발생하며 각 경우에서 기각된다는 것을 보여줍니다. 이러한 분석을 통해 제안하는 방법은 DSBi 데이터셋에 대해 약 0.03초/이미지가 소요되며, 기존 점자 인식 기술을 능가하며, 실시간 응용을 위한 최첨단 성능을 제공한다는 결론을 내릴 수 있다.

IV. 결론 및 향후 연구

본 논문에서는 경량 컨벌루션 신경망을 이용한 자동 점자 이미지 문자 인식 알고리즘을 제안한다. 전처리 단계에서는 물체의 진전 방향, 즉 점자 영상의 점자점 패턴을 평가하여 오정렬 문제를 해결하기 위해 주성분 분석 전략을 적용하였다. 또한 수학, 적응 및 기하 연산자를 사용하여 점자 이미지 기능을 개선합니다.

그 후, 제안된 경량 컨벌루션 뉴럴과 IRB를 사용하여 문자 인식을 수행한다. 최첨단 방법과 비교한 결과 설계된

이 방법은 더 나은 정확도(즉, 영어 및 중국어 점자 데이터 세트의 경우 각각 95.2% 및 98.3%)로 빠른 문자 인식을 수행하고 타베이션당 입력 이미지 노이즈에 강합니다. 제안된 모델의 시간 복잡도도 최신 방법과 비교한다. 우리의 모델은 계산 시간이 적고 점자 문자 인식을 빠르게 수행합니다(예: 영어 및 중국어 점자 이미지의 경우 각각 0.01초 및 0.03초). 최첨단 문자 인식 정확도와 속도는 제안이 실시간 응용 프로그램에 적합할 수 있음을 증명합니다. 앞으로 이 방법을 확장하여 더 큰 점자 데이터 세트에서 인식을 수행할 수 있습니다. 향후 작업에서는 이미지 소스의 차이로 인해 점자 문자 인식 시스템에 대한 점자 이미지 품질의 영향을 조사할 수도 있습니다.

참조

[1] S. Shokat, R. Riaz, SS Rizvi, K. Khan, F. Riaz, SJ Kwon, "점자 텍스트 변환 방법의 분석 및 평가" Mobile Inf. Syst., vol. 2020, pp. 1–14, 2020년 7월, doi: 10.1155/2020/3461651.

[2] A. Mousa, H. Hiary, R. Alomari 및 L. Alnemer, "스마트 점자 시스템 인식기", Int. J. 컴퓨트 과학. 문제, 권. 10, 아니. 6, p. 52, 2013.

[3] Y. LeCun, Y. Bengio 및 G. Hinton, "딥 러닝" Nature, vol. 521, 아니요. 7553, p. 436, 2015년 2월.

[4] T. Patil, S. Pandey, K. Visrani, "A review on basic deep learning technologies and applications", 데이터 과학 및 지능형 애플리케이션 (데이터 엔지니어링 및 통신 기술 강의 노트). 2021.

[5] R. Aggarwal, V. Sounderajah, G. Martin, DSW Ting, A. Karthikesalingam, D. King, H. Ashrafian 및 A. Darzi, "의료 영상에서 심층 학습의 진단 정확도: 체계적 검토 및 메타 분석," NPJ Digit. Med., 권. 4, 아니요. 1, pp. 1–23, 2021년 12월, doi: 10.1038/s41746-021-00438-z.

[6] T. Kausar, MA Ashraf, A. Kausar 및 I. Riaz, "유방암 유형 분류를 위한 확산 신경망 기반 접근법", Proc. 국제 부르반 회의 신경 과학. 기술. (IBCAST), 2021년 1월, pp. 407–413, doi: 10.1109/IBCAST51254.2021.9393249.

[7] X. Wang, Y. Yan, P. Tang, X. Bai, W. Liu, "다중 인스턴스 신경망 재방문" Pattern Recognit., vol. 74, pp. 15–24, 2018년 2월, doi: 10.1016/j.patcog.2017.08.026.

[8] P. Wang, E. Fan, and P. Wang, "전통적인 기계 학습과 딥 러닝을 기반으로 한 이미지 분류 알고리즘의 비교 분석" Pattern Recognit. Lett., vol. 141, pp. 61–67, 2021년 1월, doi: 10.1016/j.patrec.2020.07.042.

[9] Y. Yuan, C. Wang 및 Z. Jiang, "스펙트럼-공간 초분광 이미지 분류를 위한 프록시 기반 딥 러닝 프레임워크: 효율적이고 견고함" IEEE Trans. 지리학. Remote Sens., vol. 60, pp. 1–15, 2021, doi: 10.1109/TGRS.2021.3054008.

[10] Y. Liu, P. Sun, N. Wergeles 및 Y. Shang, "작은 물체 감지를 위한 심층 학습 방법에 대한 조사 및 성능 평가" Expert Syst. Appl., vol. 172, 2021년 6월, Art. 아니요. 114602, doi: 10.1016/j.eswa.2021.114602.



- [11] AW Kabani 및 MR El-Sakka, Proc. 국제 회의 이미지 향문. *Recognit.*, 2016, pp. 358–366, doi: [10.1007/978-3-319-41501-7\\_41](https://doi.org/10.1007/978-3-319-41501-7_41).
- [12] B. Wu, T. Kausar, Q. Xiao, M. Wang, W. Wang, B. Fan, D. Sun, "FF CNN: 유방암 조직학에서 유사분열 검출을 위한 효율적인 심층 신경망 이미지", *Proc. 안누. 회의 중간 Image Understanding Anal.*, 2017, pp. 249–260, doi: [10.1007/978-3-319-60964-5\\_22](https://doi.org/10.1007/978-3-319-60964-5_22).
- [13] J. Ji, X. Lu, M. Luo, M. Yin, Q. Miao 및 X. Liu, "시맨틱 분할을 위한 병렬 완전 컨볼루션 네트워크" *IEEE Access*, vol. 9, pp. 673–682, 2021, doi: [10.1109/ACCESS.2020.3042254](https://doi.org/10.1109/ACCESS.2020.3042254).
- [14] A. Ouahabi 및 A. Taleb-Ahmed, "실시간 시맨틱 분할을 위한 딥 러닝: 초음파 이미지에 적용", *Pattern Recognit. Lett.*, vol. 144, pp. 27–34, 2021년 4월, doi: [10.1016/j.patrec.2021.01.010](https://doi.org/10.1016/j.patrec.2021.01.010).
- [15] N. Alalwan, A. Abozeid, AA ElHabshy 및 A. Alzahrani, "의료 영상 의미론적 세분화를 위한 효율적인 3D 딥 러닝 모델", *Alexandria Eng. J.*, vol. 60, 아니. 1, pp. 1231–1239, 2021년 2월, doi: [10.1016/j.aej.2020.10.046](https://doi.org/10.1016/j.aej.2020.10.046).
- [16] Z. Song, "다기능 딥러닝 기반 영어 음성 인식" *Computing*, vol. 102, 아니. 3, pp. 663–682, 2020년 3월, doi: [10.1007/s00607-019-00753-0](https://doi.org/10.1007/s00607-019-00753-0).
- [17] Y. Jin, B. Wen, Z. Gu, X. Jiang, X. Shu, Z. Zeng, Y. Zhang, Z. Guo, Y. Chen, T. Zheng, Y. Yue, H. Zhang, H. Ding, "딥 러닝 지원 MXene 기반 인공 목: 소리 감지 및 음성 인식을 향하여" *Adv. 얼마. 기술*, 권. 2020년 7월 9일, Art. 아니요. 2000262, doi: [10.1002/admt.202000262](https://doi.org/10.1002/admt.202000262).
- [18] T. Kausar, M. Wang, MA Ashraf 및 A. Kausar, "SmallMitosis: 유방 조직병리학 이미지에서 작은 크기 유사분열 세포 검출" *IEEE Access*, vol. 9, pp. 905–922, 2021, doi: [10.1109/ACCESS.2020.3044625](https://doi.org/10.1109/ACCESS.2020.3044625).
- [19] T. Kausar, M. Wang, M. Idrees 및 Y. Lu, "HWCNN: Haar 웨이블릿 분해 이미지 기반 합성곱 신경망을 이용한 유방 조직병리학의 다중 클래스 인식", *Biocybern. 바이오메드. Eng.*, vol. 39, 아니. 4, pp. 967–982, 2019년 10월, doi: [10.1016/j.bbe.2019.09.003](https://doi.org/10.1016/j.bbe.2019.09.003).
- [20] Y. Lu, M. Wang, W. Wu, Q. Zhang, Y. Han, T. Kausar, S. Chen, M. Liu 및 B. Wang, "단수 기반 엔트로피 기반 패턴 학습 생리학적 신호 평가를 위한 스펙트럼 분석 구성요소", *Complexity*, vol. 2020, pp. 1–17, 2020년 1월, doi: [10.1155/2020/4625218](https://doi.org/10.1155/2020/4625218).
- [21] M. Bakator 및 D. Radosav, "딥 러닝 및 의학적 진단: 문헌 검토", *Multimodal Technol. 인터랙트*, vol. 2, 아니요. 3, p. 47, 2018년 8월, doi: [10.3390/mti2030047](https://doi.org/10.3390/mti2030047).
- [22] VV Murthy, M. Hanumanthappa 및 S. Vijayanand, "캐니 에지 검출기를 사용한 점자 셀 분할 및 원치 않는 도트 제거", *Advances in Artificial Intelligence and Data Engineering*, 2021, 도이: [10.1007/978-981-15-3514-7\\_7](https://doi.org/10.1007/978-981-15-3514-7_7).
- [23] IG Ovodov, "광학 점자 인식에 적용되는 존지도 머신러닝 효율 개선을 위한 시맨틱 기반 주석 강화 알고리즘", *Proc. IEEE 회의 러시아어 젊은 해상도. 선택된. 전자. 공학 (ElConRus)*, 2021년 1월, pp. 2190–2194, doi: [10.1109/ElConRus51938.2021.9396534](https://doi.org/10.1109/ElConRus51938.2021.9396534).
- [24] S. Alufaisan, W. Albur, S. Alsedrah 및 G. Latif, "컨볼루션 신경망을 사용한 아라비아 점자 숫자 인식", *Proc. 국제 회의 Commun., Comput. 전자. Syst.*, 2021, p. 87, doi: [10.1007/978-981-33-4909-4\\_7](https://doi.org/10.1007/978-981-33-4909-4_7).
- [25] RF Turkson, F. Yan, MKA Ali 및 J. Hu, "불꽃 점화 엔진 보정의 인공 신경망 응용: 개요" *Eng. 과학. Technol.*, Int. J., vol. 19, 아니. 3, pp. 1346–1359, 2016년 9월, doi: [10.1016/j.jestech.2016.03.003](https://doi.org/10.1016/j.jestech.2016.03.003).
- [26] S. T 및 V. Udayashankara, "엔보싱된 점자 문자의 광학 인식을 위한 소프트웨어 알고리즘에 대한 리뷰" *Int. J. 컴퓨트 Appl.*, vol. 81, 아니요. 3, pp. 25–35, 2013년 11월, doi: [10.5120/13993-2015](https://doi.org/10.5120/13993-2015).
- [27] H. Burton, AZ Snyder, TE Conturo, E. Akbudak, JM Ollinger 및 ME Raichle, "초기 맵핑과 후기 맵핑의 적용 변화: 점자 판독에 대한 fMRI 연구", *J. Neurophysiol.*, vol. 87, 아니. 1, pp. 589–607, 2002년 1월, doi: [10.1152/jn.00285.2001](https://doi.org/10.1152/jn.00285.2001).
- [28] IG Ovodov, "객체 감지 CNN을 이용한 광학 점자 인식" 2020, arXiv:2012.12412.
- [29] B.-M. Hsu, "사각장애인과 비사각장애인 간 비대칭 의사소통 감소를 위한 점자 인식" *Symmetry*, vol. 12, 아니. 7, p. 1069, 2020년 6월, doi: [10.3390/SYM12071069](https://doi.org/10.3390/SYM12071069).
- [30] J. Mao, J. Zhu, X. Wang, H. Liu 및 Y. Qian, "제한된 교육 데이터를 사용한 중국어 점자의 음성 합성", *Proc. IEEE Int. 회의 멀티미디어 엑스포(ICME)*, 2021년 7월, pp. 1–6, doi: [10.1109/icme51207.2021.9428160](https://doi.org/10.1109/icme51207.2021.9428160).
- [31] P. Kaur, S. Ramu, S. Panchakshari 및 N. Krupa, "이미지 및 음성 처리를 사용하여 음성으로 힌디어 점자의 변환" *Proc. IEEE 7th Uttar Pradesh Section Int. 회의 일렉트릭, 일렉트론. 컴퓨팅 공학 (UPCON)*, 2020년 11월, pp. 1–6, doi: [10.1109/UPCON50219.2020.9376566](https://doi.org/10.1109/UPCON50219.2020.9376566).
- [32] AA Choudhury, R. Saha, SZH Shoumo, SR Tulon, J. Uddin 및 MK Rahman, "YOLO 알고리즘을 사용하여 점자를 나타내는 효율적인 방법", *Proc. 합동 7th Int. 회의 알파니다., 전자. 비스. (ICIEV)*, 2019년 6월, 379–383쪽, doi: [10.1109/ICIEV.2018.8641038](https://doi.org/10.1109/ICIEV.2018.8641038).
- [33] M. Jiang, X. Zhu, G. Gielen, E. Drábek, Y. Xia, G. Tan 및 T. Bao, "중국어 번역을 인쇄하기 위한 점자" *Inf. 소프트 기술*, 권. 44, pp. 91–100, 2002년 2월, doi: [10.1016/S0950-5849\(01\)00220-8](https://doi.org/10.1016/S0950-5849(01)00220-8).
- [34] W. 리저츠. (2020). 음악 점자 교육학: 설명, 점자, 음악 학습 이론 및 라반의 교차점. [온라인]. 이용 가능: <https://researchspace.auckland.ac.nz/handle/2292/51655>.
- [35] S. Shokat, R. Riaz, SS Rizvi, AM Abbasi, AA Abbasi, SJ Kwon, "위치 자유 터치스크린 기반 점자 입력 방식을 이용한 문자 예측을 위한 딥러닝 기반" *Hum.-centric 컴퓨팅 Inf. 과학*, 권. 10, 아니. 1, pp. 1–24, 2020년 12월, doi: [10.1186/s13673-020-00246-6](https://doi.org/10.1186/s13673-020-00246-6).
- [36] R. Li, H. Liu, X. Wang, J. Xu, Y. Qian, "보조 학습 전략을 사용한 의미 분할 네트워크 기반 광학 점자 인식" *Proc. IEEE/CVF 회의 컴퓨팅 비스. 패턴 인식. 워크샵*, 2020년 6월, pp. 554–555, doi: [10.1109/CVPRW50498.2020.00285](https://doi.org/10.1109/CVPRW50498.2020.00285).
- [37] O. Ronneberger, P. Fischer 및 T. Brox, "U-Net: 생의학 이미지 분할을 위한 컨볼루션 네트워크", *Proc. 국제 회의 중간 이미지 컴퓨팅. Comput.-아사스트. Intervent.*, 2015, pp. 234–241, doi: [10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- [38] R. Li, H. Liu, X. Wang, Y. Qian, "양면 점자 이미지에 대한 2단계 학습 기반의 효과적인 광학 점자 인식", *Proc. 퍼시픽람인터네셔널 회의 아티프. Intell.*, 2019, pp. 150–163, doi: [10.1007/978-3-030-29894-4\\_12](https://doi.org/10.1007/978-3-030-29894-4_12).
- [39] T.-Y. Lin, P. Goyal, R. Girshick, K. He 및 P. Dollar, "고밀도 객체 감지를 위한 초점 손실" *IEEE Trans. 패턴 형문. 마하. Intell.*, vol. 42, 아니. 2, pp. 318–327, 2020년 2월, doi: [10.1109/TPAMI.2018.2858826](https://doi.org/10.1109/TPAMI.2018.2858826).
- [40] R. Li, H. Liu, X. Wang 및 Y. Qian, "DSBI: 점자 점 감지를 위한 양면 점자 이미지 데이터 세트 및 알고리즘 평가", *Proc. 2nd Int. 회의 비디오 이미지 프로세스.*, 2018, pp. 65–69, doi: [10.1145/3301506.3301532](https://doi.org/10.1145/3301506.3301532).
- [41] 제목 없음. 접속일: 2020년 8월 20일. [온라인]. 이용 가능: <https://www.kaggle.com/shanks0465/braille-character-dataset/> [42] A. Krizhevsky, I. Sutskever 및 GE Hinton, "심층 컨볼루션 신경망을 사용한 ImageNet 분류" *Commun. ACM*, vol. 60, 아니. 6, pp. 84–90, 2017년 5월, doi: [10.1145/3065386](https://doi.org/10.1145/3065386).
- [43] K. Simonyan 및 A. Zisserman, "Very deep convolutional network for 대규모 이미지 인식", *Tech. 의원*, 2015.
- [44] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens 및 Z. Wojna, "컴퓨터 비전을 위한 초기 아키텍처 재검토", *Proc. IEEE 회의 컴퓨팅 비스. 패턴 인식. (CVPR)*, 2016년 6월, pp. 2818–2826, doi: [10.1109/CVPR.2016.308](https://doi.org/10.1109/CVPR.2016.308).
- [45] K. He, X. Zhang, S. Ren, J. Sun, "이미지 인식을 위한 심층 잔차 학습" *Proc. IEEE 회의 컴퓨팅 비스. 패턴 인식. (CVPR)*, 2016년 6월, pp. 770–778, doi: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90).
- [46] G. Huang, Z. Liu, L. Van Der Maaten 및 KQ Weinberger, *Proc. IEEE 회의 컴. 비스. 패턴 인식. (CVPR)*, 2017년 7월, pp. 4700–4708, doi: [10.1109/CVPR.2017.243](https://doi.org/10.1109/CVPR.2017.243).
- [47] M. Lin, Q. Chen 및 S. Yan, "네트워크의 네트워크" *Tech. 의원*, 2014.
- [48] G. Yadav, S. Maheshwari 및 A. Agarwal, "실시간 비디오 시스템을 위한 대비 제한적 적응 히스로그램 등화 기반 향상", *Proc. 국제 회의 고급 계산. Commun. 알라다. (ICACCI)*, 2014년 9월, pp. 2392–2397, doi: [10.1109/ICACCI.2014.6968381](https://doi.org/10.1109/ICACCI.2014.6968381).
- [49] R. Kreslin, PM Calvo, LG Corzo 및 P. Peer, "Delaunay 삼각분할에 기반한 선형 색채 적응 변환" *Math. 문제 과학*, vol. 2014, pp. 1–9, 2014년 1월, doi: [10.1155/2014/760123](https://doi.org/10.1155/2014/760123).
- [50] AK Singh 및 BK Singh, "디지털 이미지 처리에서 인간 생체 인식의 응용" *Int. J. Innov. 과학. 해상도 기술*, 권. 5, 아니요. 7, pp. 1273–1276, 2020년 8월, doi: [10.38124/ijisrt20jul748](https://doi.org/10.38124/ijisrt20jul748).
- [51] HZU Rehman 및 S. Lee, "주성분 분석을 이용한 자동 이미지 정렬" *IEEE Access*, vol. 6, pp. 72063–72072, 2018, doi: [10.1109/ACCESS.2018.2882070](https://doi.org/10.1109/ACCESS.2018.2882070).
- [52] TZ Xiang, GS Xia, X. Bai, L. Zhang, "전역 유사성 제약 조건이 있는 선 안내 로컬 뒤틀럼에 의한 이미지 스타킹" *Pattern Recognit.*, vol. 83, pp. 481–497, 2018년 11월, doi: [10.1016/j.patcog.2018.06.013](https://doi.org/10.1016/j.patcog.2018.06.013).

- [53] G. Zhang, L. Pan, J. Chen, Y. Gong 및 F. Liu, "이미지 면의 공동 정렬" IEEE Access, vol. 8, pp. 114884–114891, 2020, doi: [10.1109/ACCESS.2020.3003332](https://doi.org/10.1109/ACCESS.2020.3003332).
- [54] HT Likassa, "아핀 변환, Frobenius 및 L2,1 규범 을 통한 관절 이미지 정렬 및 복구를 위한 새로운 강력한 주성분 분석" Int. J. 수학. 과학, 권. 2020, pp. 1–9, 2020년 4월, doi: [10.1155/2020/8136384](https://doi.org/10.1155/2020/8136384).
- [55] H. Huang, J. Zhang, J. Zhang, J. Xu 및 Q. Wu, "몇 번의 촬영 미세 입자 이미지 분류를 위한 낮은 순위 쌍 방향 정렬 이중선형 네트워크" IEEE Trans. 멀티미디어, vol. 23, pp. 1666–1680, 2021, doi: [10.1109/TMM.2020.3001510](https://doi.org/10.1109/TMM.2020.3001510).
- [56] T. Madeira, M. Oliveira 및 P. Dias, "가점 마커를 사용한 RGB-D 이미지 정렬 향상" Sensors, vol. 20, 아니. 5, p. 1497, 2020년 3월, doi: [10.3390/s20051497](https://doi.org/10.3390/s20051497).
- [57] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov 및 LC Chen, "MobileNetV2: 반전 잔차 및 선택 병목 현상", Proc. IEEE 회의 컴퓨팅 비스. 패턴 인식. (CVPR), 2018년 6월, pp. 4510–4520, doi: [10.1109/CVPR.2018.00474](https://doi.org/10.1109/CVPR.2018.00474).
- [58] DSK Jadwa, "Wiener 필터 기반 의료 영상 노이즈 제거" Int. J. Sci. 공학 Appl., vol. 7, 아니오. 9, pp. 318–323, 2018년 9월, doi: [10.7753/ijsea0709.1014](https://doi.org/10.7753/ijsea0709.1014).
- [59] N. Otsu, "그레이 레벨 히스토그램의 임계값 선택 방법" Tech. 의원, 1979.
- [60] 형태학적 이미지 분석: 원리 및 응용, Sens. Rev., 2000, doi: [10.1108/sr.2000.08720cae.001](https://doi.org/10.1108/sr.2000.08720cae.001).
- [61] K. He 및 J. Sun, "Convolutional Neural Networks at Constrained Time Cost" Proc. IEEE 회의 컴퓨팅 비스. Pattern Recognit., 2015년 6월, pp. 5353–5360, doi: [10.1109/CVPR.2015.7299173](https://doi.org/10.1109/CVPR.2015.7299173).
- [62] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard 및 M. Kudlur, "TensorFlow: 대규모 기계 학습을 위한 시스템" in Proc. 2nd USENIX Symp. 운영 체제 디자인 구현. (OSDI), 2016, pp. 265–283.
- [63] F. Chollet, "Keras," J. Chem. Inf. 모델., 2013.
- [64] L. Smith 및 Y. Gal, "대적 사례 감지를 위한 불확실성 측정 이해" Tech. 의원, 2018.
- [65] A. Hore 및 D. Ziou, "이미지 품질 메트릭: PSNR 대 SSIM", Proc. 20th Int. 회의 Pattern Recognit., 2010년 8월, pp. 2366–2369, doi: [10.1109/ICPR.2010.579](https://doi.org/10.1109/ICPR.2010.579).
- [66] AD Kiureghian 및 O. Ditlevsen, "우행 또는 인식? 상관없나요?" 구조체. Saf., vol. 31, 아니. 2, pp. 105–112, 2009년 3월, doi: [10.1016/j.strusafe.2008.06.020](https://doi.org/10.1016/j.strusafe.2008.06.020).
- [67] T. Makkar, Y. Kumar, AK Dubey, A. Rocha 및 A. Goyal, '손으로 쓴 숫자를 인식하는 KNN 및 CNN의 시간 복잡도 분석', Proc. 4위 회의 이미지정보 프로세스. (ICIIP), 2017년 12월, pp. 1–6, doi: [10.1109/ICIIP.2017.8313707](https://doi.org/10.1109/ICIIP.2017.8313707).
- [68] C. Malon, E. Brachtel, E. Cosatto, HP Graf, A. Kurata, M. Kuroda, JS Meyer, A. Saito, S. Wu, and Y. Yagi, "유사분열 형태 인식: 합의 병리학자 및 전산화된 검출기," 향문. 셀. Pathol., vol. 35, pp. 97–100, 2012년 1월, doi: [10.3233/ACP-2011-0029](https://doi.org/10.3233/ACP-2011-0029).



SAJJAD MANZOOR 는 2016년에 한국 한양 대학교에서 전자 시스템 공학 학위를 받았습니다. 그는 현재 미르푸르 과학 기술 대학교(MUST)의 미르푸르 공과대학 조교수입니다. 그의 연구 관심사는 로봇 공학, 비전 및 표적 추적을 포함합니다.



ADEEBA KAUSAR 는 B.Sc. 2017년 파키스탄 Azad Jammu 및 Kashmir 대학교에서 컴퓨터 시스템 공학 학위를 취득했으며 M.Sc. 2020년 파키스탄 타실라 공과대학에서 컴퓨터 시스템 공학 학위를 취득했습니다. 그녀는 현재 파키스탄 나로왈 대학교 컴퓨터 과학과 강사로 재직 중입니다.

그녀의 연구 관심 분야는 컴퓨터 비전, 이미지 처리, 딥 러닝 기술, 차세대 무선 통신 시스템 및 무선 센서 네트워크입니다.



YUN LU 는 1985년 중국 후난성 형양에서 태어났습니다. 2009년 후난성 상탄 대학교에서 마이크로전자 공학 학사 학위를, 2011년 중국 쑤원 대학교에서 광학 공학 석사 학위를 취득했습니다. 박사 2020년 중국 선전 심전에 있는 Harbin Institute of Technology에서 마이크로 전자 공학 및 고체 전자 공학 학위를 취득했습니다.

2011년부터 2014년까지 그는 중국 과학 아카데미의 Shen zhen Institutes of Advanced Technology에서 생물 의학 응용 분야를 위한 혼합 신호 프론트 엔드 회로 설계 작업을 했습니다. 2014년부터 2017년까지 Launch Tech Company Ltd.의 선임 엔지니어이기도 했습니다. 2020년 8월부터 중국 광둥성 Huizhou University에서 부교수로 재직하고 있습니다. 그의 현재 연구 관심 분야는 신경 과학에서 공학에 이르는 인지 컴퓨팅, 뇌-기계 인터페이스를 위한 혁신적인 방법 및 기계 학습입니다.



MUHAMMAD WASIF 는 B.Sc. 2006년 파키스탄 Riphah University에서 전기 공학 학위를 취득했으며 M.Sc. 메카트로닉스 및 Ph.D. 2008년과 2016년에 각각 영국 King's College London에서 로봇공학 학위를 받았습니다. 현재 파키스탄 구지라트 대학교 전기공학과 조교수로 재직 중입니다. 그는 Gujrat 대학교 전기 공학과 지능형 시스템 연구실의 그룹 리더입니다. 그의 연구 관심 분야는 로봇 공학, 지능형 제어, 능동 비전, 머신 비전 및 머신 러닝입니다.



M. ADNAN ASHRAF 는 B.Sc. 및 M.Sc. 2009년과 2020년에 각각 파키스탄의 Mirpur University of Science and Technology에서 전기 공학 학위를 취득했으며 현재 박사 과정을 밟고 있습니다. 통제 정도.

그는 2011년부터 2012년까지 파키스탄 왈라에 있는 고전압 및 단락 회로 실험실에서 수습 엔지니어로 일하기 시작했습니다. 그는 Transfopower Pvt.의 설계 엔지니어였습니다. Ltd., 파키스탄, 2012년부터 2014년까지. 그는 2016년에 Mirpur University of Science and Technology에서 강사로 합류했습니다.



TASLEEM KAUSAR 는 1988년 파키스탄에서 태어났습니다. 2010년 파키스탄의 Azad Jammu 및 Kashmir 대학교에서 전기 공학 학위를 취득한 M.Sc. 2013년 파키스탄 타실라에 있는 공과대학에서 전자 공학 학위를 취득하고 박사 학위를 받았습니다. 2020년 중국 하얼빈 연구소에서 전자 공학 학위를 받았습니다. 2013년부터 파키스탄 미르푸르 과학 기술 대학교에서 강사로 일하고 있습니다. 그녀의 연구 관심 분야는 딥 러닝 기술을 사용한 의료 영상 분석입니다.