

Facial Emotion Recognition을 활용한 군 간부 복무적응도 검사 개선방안 연구

조성진¹ · 김수환^{2*}

¹육군3사관학교 국방시스템학과, ²국방대학교 국방과학학부

A Study on the Improvement of Adaptability Test for Military Executive Using Facial Emotion Recognition

Sungjin Cho¹ · Suhwan Kim²

¹Department of Defense System Engineering, Korea Army Academy at YeongCheon

²Department of Defense Science, Korea National Defense University

In this study, a test model was created and studied using the ConvNeXt technique, a Facial emotion recording technique, and Tensorflow-based TFRecord data to improve the military service adaptation test method. This study through image classification techniques was first applied to military service adaptation tests to compensate for the problem of possible distortion of the test results of the subject, which is a vulnerability of the self-report psychological test currently being conducted. In addition, this study was conducted to reduce learning time and improve accuracy compared to the general model of facial emotion recognition analysis through image classification. In addition, the model created was tested on actual subjects and a way to apply it to military service adaptation tests was sought and significance was derived. This prevents the subject from intentionally distorting the test results for the test, and gives valuable insight by confirming the sincerity of the test response and a deep understanding of the emotional response of the psychological test.

Keywords: ConvNeXt, TFRecord, Facial Emotion Recognition, Military Service Adaptation Tests

1. 서론

군은 인격 특성 검사를 기반으로 사고와 사건을 예방하고, 사회 부적응 및 적응 불능의 특성을 사전에 식별하고 관리하기 위해 2015년에 국방연구원(KIDA)에서 개발한 복무적응도 검사 2.0를 사용하고 있다(Kim, 2018). 이는 군에서 지휘관, 지휘자가 장병들의 심리상태와 복무 간 어려운 점을 파악할 수 있는 가장 기본적인 수단이자 중요한 도구이다. 또한, 응답자의 심리적 상태를 종합적으로 평가하고, 이 결과를 지휘자들이 참고할 수 있는 자료로 제공한다. 특히 현역 복무 부적합 처리를 위해 사용되는 기초자료로 제공되기도 하는 복무적응도검

사 결과는 그만큼 굉장히 중요하다.

<Figure 1>과 같이 자대 내에서 주기적으로 실시하는 해당 검사를 통해 정신적인 문제나 복무 부적응 여부를 초기에 식별한다. 이후 복무 부적응을 호소하는 장병에게는 병역 처분을 변경하고 복무 적응을 지원한다(Park, 2022).

최근 군대 내에서는 간부들이 민원을 피하기 위해 군 생활을 하기 싫어 현역부적합심사를 받게 해달라는 장병들의 요구를 그대로 들어주거나 일부 간부들이 고의로 위험행동을 해 소위 '관심간부'로 분류받고 전역되는 경우가 적지 않아 철저한 조사가 요구된다는 지적이 나오고 있다. 군별로는 해군이 96.8%로 가장 높았고 육군(96%), 해병대(95.4%), 공군(95.2%) 순으로 비

* 연락저자 : 김수환 교수, 충청남도 논산시 양촌면 황산별로 1040 국방대학교 국방관리대학원, Tel : 041-831-5372,
E-mail : ksoohwan@gmail.com

2023년 10월 26일 접수; 2023년 12월 18일 수정본 접수; 2024년 3월 11일 게재 확정.

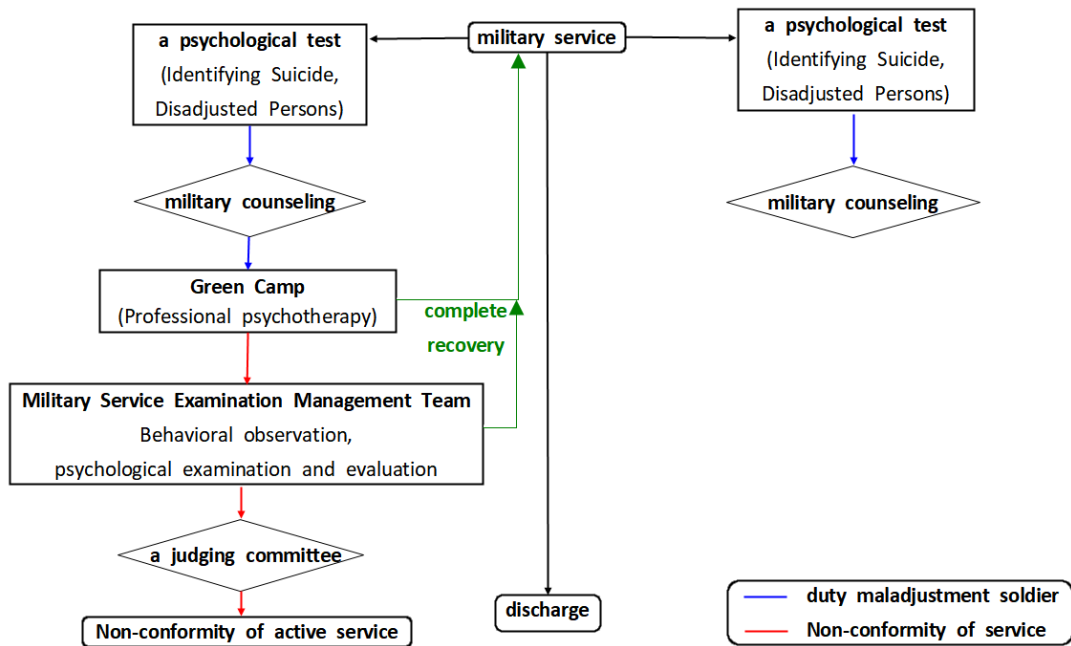


Figure 1. Support System for Service Adaptation to Military Service during Service

스한 수준이었다. 통계자료를 통해 알 수 있듯이 현역부적합심사에 오르면 제대할 확률이 96%나 될 만큼 매우 높은 것을 확인할 수 있다(Kwon, 2023). 따라서 구분별한 현역복무부적합 처리를 방지하기 위해 신뢰성과 정확성이 높은 검사가 매우 중요하다. 또한, 인구절벽으로 입대 인원 자체도 부족한 상황에서 중도이탈 인원까지 구분별하게 발생한다면 매우 심각한 병력 부족과 군사력 저하 문제를 겪을 것으로 예상된다.

하지만 현재 현역복무부적합 처리를 위해 기초자료로 제공되는 신인성검사 체계 내의 복무적응도 검사는 부적합자에 대한 선별을 제대로 할 수 없다는 맹점이 식별되고 있다. 피검자가 본인의 현재 상태를 스스로 진단 및 파악하고 이에 대한 결과를 각 질문마다 본인 스스로 선택하는 자기 보고식 응답방법의 검사 방식은 피검자가 의도적으로 결과를 긍정·부정왜곡할 수 있다는 취약점을 갖고 있다. 이러한 약점을 활용해서 복무기피를 위해 현역복무부적합의 조건을 갖추고자 하는 인원들이 이를 악용한다면 더 큰 문제점을 초래할 수 있다. 따라서, 본 논문에서는 심리검사 왜곡정도 파악을 위해 얼굴표정 이미지 데이터를 활용하여 딥러닝 기법이 적용된 실시간 얼굴 표정 감정분석 모델을 제시하고자 한다. 또한, 실험을 통해 해당 방법의 효과를 검증하고자 한다. 제시하는 모델은 학습시간을 단축시키기 위해 Tensorflow기반의 TFRecord 데이터를 이용하였으며 모델의 성능을 향상시키고자 ConvNeXt기법을 모델에 적용하였다.

본 논문의 구성은 다음과 같다. 제2장에서 관련연구에 대해 소개하고, 제3장에서 이론적배경에 대해 설명하며, 제4장에서는 제시된 연구모형을 통해 실험을 진행하고 분석결과를 도출한다. 제5장에서는 결론(연구 결과와 향후 연구방향)으로 구성하였다.

2. 관련연구

많은 연구자들이 앞서 제기한 복무적응도 검사의 취약점을 보완하기 위해 다양한 방법을 활용하여 연구하고 있다. Jung (2017)은 피문(지문)과 대뇌사고의 연관성에 대한 연구자료 분석과 군 부대 부적응 병력(신병, 관심병사, 도움 및 배려병사)에 대한 인성검사, 상담기법의 피문학 적용 방안 연구를 하였다. Nam(2019)은 복무부적응 용사들의 행동 양식 및 증상들을 온톨로지로 설계하여 사고 우려자 식별을 위한 보조 도구로 활용할 수 있도록 온톨로지 구현프로세스를 제안하였다. 복무부적응 용사 식별을 위한 객관적인 기준 제공 및 정보 공유를 위해 온톨로지 구현을 연구하였다. Kim(2021)은 과학적 기법인 생체인식 지표 중 뇌파, 맥파지표를 이용하여 신인성검사 2.0과 뇌파/맥파지표의 관계성을 확인하였다. Park(2022)는 사고 예방에 객관적인 기준을 제공하고자 복무적응도 검사 문항 개선에 대한 분석을 통해 세대 특성에 따른 검사척도 변화를 파악하고, 검사결과를 활용하여 복무 부적응 용사를 예측할 수 있는 기법 간의 분석을 하였다. 하지만 위에서 제시한 연구들은 피검자가 의도적으로 결과를 조작할 수 있는 자기보고식 심리검사 방식의 문제점을 해결하기 위해 다양한 연구를 진행하였지만 이와 같은 방법이 현재 실제 검사에 적용되고 있는 연구는 찾아 볼 수 없었다. 또한, 검증을 위해 해당 방법을 이용하여 실험까지 연구된 논문은 36명의 소규모 인원으로 구성된 Kim(2022)의 뇌파/맥파지표의 관계성 확인을 위한 연구의 다른 연구에서는 확인할 수 없었다. 하지만 뇌파/맥파를 측정한다는 것은 검사장비 보유에 대한 비용 부담이 증가하는 단점을 갖고 있으며, 현재 용사 대상으로 실시하고 있는 복무적응도 검사는 모바일을 통해 시간과 장소에 구애받지 않고 할

수 있다는 측면에서 다소 실용성이 떨어진다고 판단된다. 이를 해결하기 위해 본 논문에서는 얼굴표정 감정분석 모델을 제시하고자 얼굴표정과 관련하여 진행되고 있는 선행연구들을 확인했다. Hong(2006)은 얼굴표정이 인물의 내적 상태를 반영하고 언어를 제외한 가장 강력한 복합 신호를 준다고 하였다. LaFrance(2012)는 각각의 감정은 특정 얼굴표정과 연관되어 있고, 표정 확인을 통해 어떤 감정이 그러한 표정을 유발했는지 파악할 수 있다고 하였다. Snowden(2013)은 얼굴표정을 통해 기분을 확인할 수 있으며, 특히 얼굴표정의 조그마한 변화를 통해 사람들의 생각도 분별할 수 있다고 한다. Ekman(2014)은 ‘놀라움, 두려움, 혐오, 화, 행복, 슬픔’의 6가지 감정을 드러내는 보편적인 얼굴 표정을 통해 ‘기본표정 모형’을 적용하여 감정을 인지할 수 있는 방식을 체계적으로 제시하고, 다양한 얼굴 표정 뒤에 숨은 진짜 감정을 찾아내는 방법을 연구했다. 특히 1초 미만의 짧은 순간에 나타났다 사라지는 ‘미세표정(micro-expressions)’으로 거짓과 진실을 구별해내는 법을 체계화하였으며, 수천 개의 얼굴근육을 분석해 ‘얼굴 지도(얼굴 움직임 부호화 시스템)’를 만들었다.

2. 이론적 배경

2.1 TFRecord

TFRecord 파일은 Tensorflow의 학습 데이터 등을 저장하기 위한 Binary Data format으로, 구글의 Protocol Buffer format의 Data를 파일에 Serialize하여 저장한다. <Figure 2>에서 1번은 TFRecord로 저장되기 전 개별 파일의 일련 정보다. 2번은 각각의 라벨 정보다. 3번은 각각의 이미지 정보다. 모든 정보들이 다 떨어져 있기 때문에 4번과 같이 모두 Serialize를 통해 하나로 이어 붙일 수 있다. 즉, 하나의 파일로 만드는 것이 TFRecord를 만드는 과정이다.

CSV 파일에서 숫자 또는 텍스트 데이터를 읽을 경우에는 별다른 문제점이 발생하지 않는다. 하지만 이미지 데이터를 읽을 경우 이미지는 JPEG나 PNG형태의 파일로 저장되어 있다. 이

에 대한 메타 데이터와 레이블은 별도의 파일에 저장되어 있기 때문에, 학습 데이터를 읽을 때 메타데이터나 레이블 파일 하나만 읽는 것이 아니라 이미지 파일도 별도로 읽어야 하므로 코드가 복잡해진다. 이미지를 JPG나 PNG 포맷으로 읽어서 매번 디코딩을 하게 될 경우, 성능 저하로 인해 학습 단계에서 데이터를 읽는 부분에서 많은 성능 저하가 발생한다. 따라서, TFRecord를 사용하는 것으로써 대규모 데이터를 효율적으로 학습할 수 있게 된다. TFRecord 데이터 사용의 장점을 정리하면 첫째, 효율성이다. TFRecord 형식의 데이터는 원래 데이터보다 더 적은 공간을 차지하기 때문이다. 둘째, TensorFlow는 병렬 I/O 동작으로 TFRecord 형식의 데이터를 읽을 수 있다. 이것은 GPU나 TPU 장치로 작업할 때 매우 유용하다. 셋째, TFRecord 데이터는 데이터와 그것의 메타데이터 등 필요한 모든 것을 하나의 파일에 저장할 수 있으며 사용자가 사용하고 있는 파일형식이 무엇이든 TFRecord로 변환하여 저장 및 사용할 수 있다(Woo, 2022).

2.2 ConvNeXt

ConvNeXt는 합성곱 신경망인 CNN(Convolutional Neural Network)에 속하는 모델 중 하나이다. 기존에는 복잡한 분류나 회귀 문제를 해결하기 위해 다층 퍼셉트론을 쌓는 MLP(Multi-Layer Perceptron) 기법이 개발되었다. 그러나 이미지와 같은 비정형 데이터의 경우, MLP에 적용하기 위해선 1차원 행렬로 변환해야만 한다. 기존 방법론에서는 이미지를 1차원 행렬로 변환하는 과정에서 공간 정보(Spatial Feature)가 손실될 수 있다는 문제가 있었다. 이러한 문제를 해결하기 위해 CNN은 공간 정보를 적절히 반영할 수 있다. 2012년에 발표된 AlexNet을 시작으로, VGGNet, ResNeXt, DenseNet, MobileNet, EfficientNet, RegNet 등 다양한 CNN 모델이 등장하여 발전해 왔다. 그 중 ConvNeXt는 현대적인 기법을 도입한 최신 모델로, 다음과 같은 방법으로 설계된다. 먼저, swin transformer에서 사용된 훈련 기술을 적용한다. 주로 예폭 수를 90에서 300으로 증가시킨다. 또한, Mixup, cutmix, random augment, random erasing과 같은 데이터 증강 기법을 사용하며, Adam에 가중치 감소(weight decay)를 추가한다. 둘째, Macro Design을 적용한다. <Figure 3>과 같이 Swin-Transformer는 Input을 Patch단위로 나누었으며 Hierarchical 구조를 4개의 Block단위로 분리하였는데 이에 대해 2가지를 적용시켰다.

<Figure 4>에서 기존 ResNet은 (a)와 같이 1x1 컨볼루션으로 채널을 줄인 후 3x3 컨볼루션을 진행하고 다시 1x1으로 채널을 늘리는 Bottleneck 구조를 가지고 있다. MobileNetV2에서는 FLOPs를 줄이기 위해 (b)와 같은 Inverted Bottleneck을 사용했는데, ConvNeXt에도 동일하게 적용되었다. 이는 다운샘플링 잔여 블록에 의해 많은 FLOPs가 감소되는 효과를 가질 수 있다. 또한, 커널 크기를 증가시켰다. 이전 대부분의 모델에서는 3x3 컨볼루션을 기본으로 사용했는데, 이를 7x7 형태인 큰 커

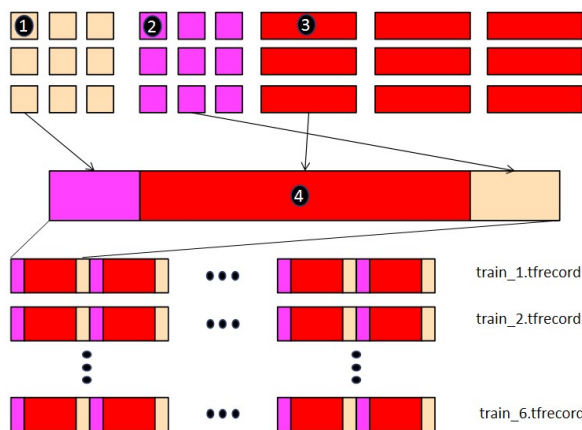


Figure 2. TFRecord Process

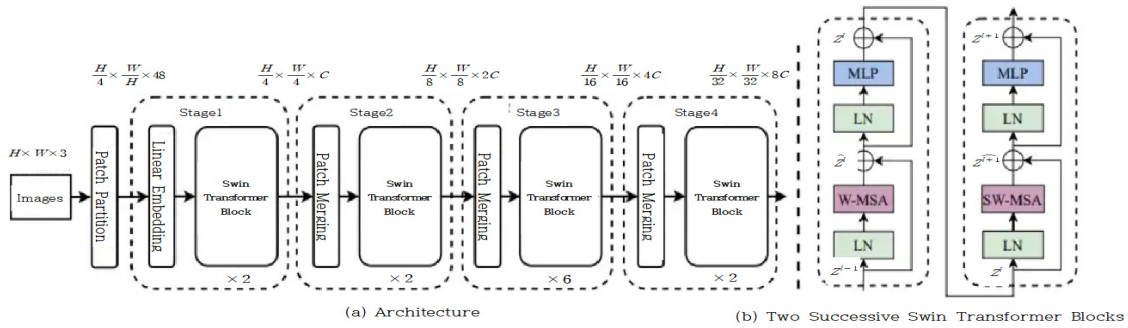


Figure 3. Swin-Transformer

널 크기로 변경했다. 더불어, 활성화 함수와 정규화 레이어 등과 같은 마이크로 스케일을 변경했다.

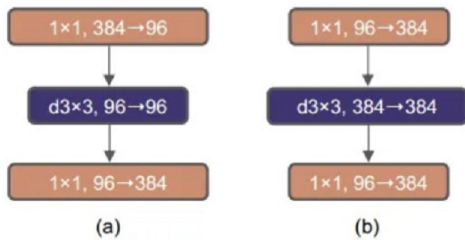


Figure 4. Inverted Bottleneck

<Figure 5>에서는 GELU를 사용을 통해 ReLU를 대체하였다. 더불어, 매우 적은 활성화 함수를 사용했는데, ConvNeXt에서는 2개의 1x1 레이어 사이에서만 활성화 함수를 적용했다. 일반적으로 ConvNet에서는 바로 뒤에 배치 정규화를 위치시키지만, ConvNeXt에서는 매우 적은 정규화를 사용한다. 이로써 배치 정규화를 줄일 수 있게 되었다. 또한, 일곱 번째로 각 스테이지의 시작 전에 다운샘플 레이어를 추가했다(Liu et al., 2020).

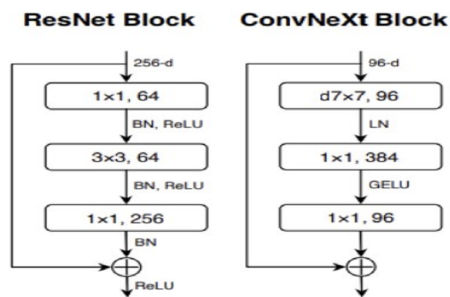


Figure 5. Fewer activation Function

3. 연구모형 및 방법

본 연구는 사람 얼굴표정 데이터를 활용하여 실시간 피검자의 얼굴표정을 분석하는 딥러닝 모델을 연구했다. 최초의 JPG 파일들을 라벨 처리한 후 TensorFlow를 기반으로 하는

TFRecord 형식으로 변환했다. 이후 해당 데이터를 ConvNeXt를 이용하여 학습시킨 뒤, OPEN-CV 기법을 활용하여 실시간으로 피검자의 얼굴표정을 탐지하고 감정을 분석했다. 연구를 위한 환경은 Windows 10 기반의 Google Colab을 사용하였고 TensorFlow 2.11.0을 활용하였다.

3.1 데이터 전처리 및 변환

Data 수집 단계에서는 7가지 감정(Happy : 행복, Surprise : 놀람, Sad : 슬픔, Fear : 두려움, Neutral : 중립, Angry : 분노, Disgusting : 당황)을 나타내는 얼굴표정 이미지 38,500장으로 구성하였다. 한국인의 얼굴표정 이미지는 한국지능정보사회진흥원에서 전문 배우로 구성되어 촬영된 이미지 35,000장(컬러)을 수집하였으며, 다문화 장병에 대비하여 외국인 얼굴표정 이미지는 Kaggle에서 3,500장(흑백) 수집하였다. Data 수집 이후에 모델 성능평가를 위해 Train Data, Validation Data, Test Data를 6:2:2로 분류하였다. 이미지 Data구성은 <Table 1>과 같다.

Kaggle에서 수집한 이미지 데이터의 해상도가 한국지능정보사회진흥원에서 수집한 이미지 데이터의 해상도보다 낮아 추후 이미지 분석에 대한 정확도를 높이기 위해 Kaggle 이미지파일의 해상도를 높이는 작업을 진행하였다. 또한, 앞서 설명한대로 JPG 형식의 데이터를 Tensorflow 기반 TFRecord 형식의 데이터로 변환해야 한다. JPG 파일형태의 사진의 경우 흑백사진과 컬러사진으로 구성되어 있기 때문에 이를 TFRecord로 변환하기 전 흑백 또는 컬러사진 둘 중의 하나의 형태로 변경해야한다. 따라서 본 연구에서는 흑백사진을 컬러사진으로 바꿔주는 코드를 통해 동일하게 모두 컬러사진의 JPG 파일로 만들었다. 이후 총 38,500장에 대해서 단 하나의 TFRecord 데이터 형태 파일로 만드는 것이 아닌 각 감정별로 TFRecord 데이터 형태 파일로 변경하였다. 이후 모델의 일반화 성능을 향상시키기 위해 데이터 증강기법을 적용하였다.

3.2 모델 구성절차

사전에 각 감정별로 JPG 파일 형태에서 TFRecord 데이터 파일 형태로 변환시킨 파일들을 Train과 Validation에 맞게 코드에

Table 1. Number of a Person's Facial Expression Images

Total		38,500 = 35,000(Korean) + 3,500(Foreigner)					
Happy		Neutral		Sad		Fear	
Korean	Foreigner	Korean	Foreigner	Korean	Foreigner	Korean	Foreigner
5,000	500	5,000	500	5,000	500	5,000	500
Surprise		Angry		Disgusted			
Korean	Foreigner	Korean	Foreigner	Korean	Foreigner		
5,000	500	5,000	500	5,000	500		

배치시켰다. 모형은 앞에 언급한대로 합성곱 신경망 모델 중 하나인 ConvNeXt 모델을 적용했다. epoch은 100으로 설정하고 학습을 진행시켰으며, 학습 과정에서 accuracy가 어느 정도 증가했다가 감소하는 지점에서 학습이 조기 종료될 수 있도록 early stop을 설정했다. 이후 학습된 모델을 바탕으로 피검자가 심리검사를 진행하면서 실시간으로 피검자의 얼굴표정을 감지하여 학습시켰던 모델을 바탕으로 감정분석을 실시할 수 있도록 OPEN-CV 기법을 적용 및 연결하였다.

3.3 모형 성능 평가

본 연구는 연구자가 제시한 TFRecord 학습속도 개선에 대

한 결과를 확인하기 위해 일반 JPG파일 가지고 CSV형태로 변환 후 학습하는 과정과 TFRecord 형태로 변환 후 학습하는 과정에 대해서 학습시간을 비교하였다. JPG의 해상도는 학습 시간과 컴퓨터 메모리 용량을 고려하고 과적합이 발생하지 않는 상태에서 추후 얼굴의 작은 변화나 세부사항을 더 잘 인식하여 이를 기반으로 감정을 잘 분류할 수 있도록 최적의 해상도 값인 1280×720으로 설정하였으며 사진 1장당 파일 크기는 1.5MB다.

또한, 모형의 성능을 확인하기 위해 검증과정을 진행했다. <Figure 7>을 통해 모형의 val값과 loss값을 확인할 수 있다.

또한, Test Data를 활용하여 학습된 모델들의 성능 비교를 실시하였다. <Table 3>은 Test Data를 활용하여 이미지 분류모

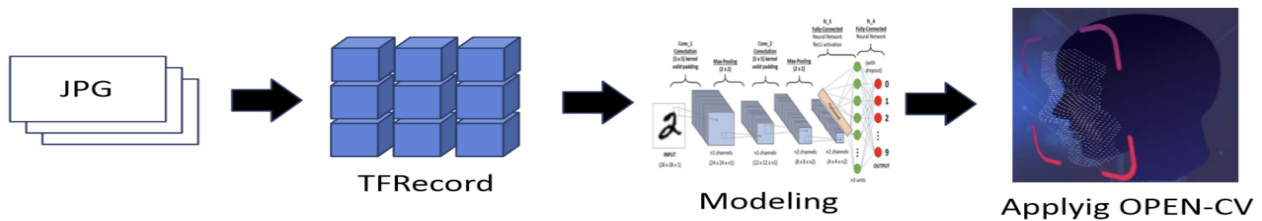


Figure 6. Modeling Process

Table 2. Comparison of Learning Speed

Total number of files	TFRecord File Size	Running time		Result
		TFRecord	JPG	
38,500 copies	6,997MB	21.4 minutes	49.2 minutes	2.3x improvement

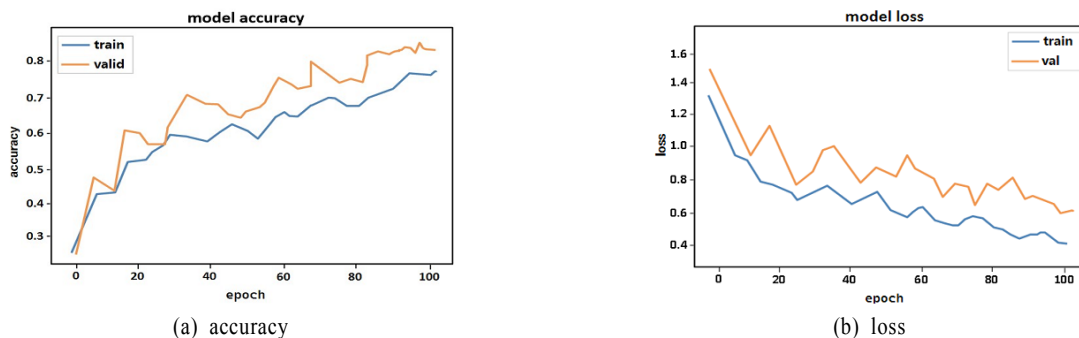


Figure 7. ConvNeXt Model Accuracy & Model Loss

델들의 총 4가지의 성능지표에 대해 Confusion Matrix로 정리하였다. 성능지표는 Accuracy(정확도), Precision(정밀도), Recall(재현율), F1 Score 이다. ConvNeXt모델이 우수한 성능을 보이는지 검증하기 위해 과거 CNN기반 중 딥러닝 모델에서 대표적으로 많이 사용하고 있는 ResNet50모델과 모바일 및 임베디드 기기와 같이 자원이 제한된 환경에서도 효율적으로 사용이 가능한 MobileNetV2 모델을 가지고 전처리부터 학습까지 동일한 조건으로 설정한 후 성능을 비교하였다. 그 결과 <Table 3>과 같이 ConvNext 모델이 가장 우수한 예측결과 값을 나타냈다.

Table 3. Image Classification Model Learning Results

Model	Accuracy	Precision	Recall	F1 Score	rank
ConvNeXt	0.8315	0.8527	0.8436	0.8344	1
ResNet50	0.7987	0.8076	0.7567	0.7899	2
MobileNetV2	0.7854	0.7982	0.7632	0.7855	3

본 연구자가 제시한 TFRecord 파일을 통해 개선된 학습 속도와 우수한 정확도를 가진 성능 좋은 모델을 활용한다면 검사 간 신뢰성 있는 결과 값을 도출할 수 있을 것으로 판단된다.

4. 검사모델을 통한 실험 진행

4.1 진행방법

실험에 참여하게 되는 피검자 선정방법은 향후 간부들에게 적용된다는 점을 가정하여 군인 신분인 간부 37명, 병사 3명으로 선정하였다. 군 간부들이 피검자의 대부분을 차지하나 병사에서 부사관으로 임관하거나 간부사관 전형을 통해 장교로 임관하는 특이케이스를 고려하여 소수의 인원인 병사 3명을 실험 대상으로 선정하였다. 질문은 자유로운 의견을 얘기할 수 있도록 비대면 방식으로 진행하고 AI가 질문을 말해주는 방식으로 검사를 실시하였다. 질문 구성은 현재 자기보고식 응답방법으로 실시하고 있는 복무적응도 검사 2.0의 질문 190

Table 4. Questions for Test

	Content
1	Tell me about your happy memories in the military
2	Tell me about a time when you were proud of yourself during your mission. Or was there a case where you were praised or awarded a commendation for making a big contribution?
3	What do you think of as a soldier? The role of a soldier?
4	Have you ever been into PC games, stocks, etc., and your service was disrupted? If you have, please be specific.
5	Have you ever experienced any stress from personal relationships, subordinates and colleagues while serving?
6	Have you ever been stressed between work?
7	Please briefly tell me what needs to be improved when looking at the military in general.
8	Have you ever thought you wanted to die because of stress from work or relationships?
9	Have you ever experienced any disruptions to your side work the next day due to drinking?
10	Are you willing to help when there are people around you asking for help?

개 중 중복적으로 묻는 문항을 분석 및 도출하여 <Table 4>와 같이 총 10개의 질문으로 구성하였다.

<Figure 8>은 피검자를 대상으로 검사 시 응답한 얼굴표정에 대해 감정을 분석하는 장면을 가정하여 인공지능으로 사람 모델을 만들어 분석모델에 적용한 예시자료다. 피검자가 검사를 진행할 때 검사 당시 화면에는 <Figure 8>과 같이 실시간 감정에 대해 분석한 수치가 나타나진 않지만 별도로 감정들을 분석하여 해당 수치들을 파일로 정리하여 결과에 활용한다. 검사를 진행할 때 본인의 감정이 실시간으로 분석되어 나오는 수치가 화면에 나타날 시 응답에 영향을 미치므로 실시간으로 감정이 분석되고 있는 수치들은 피검자의 화면에서는 나타나지 않도록 한다.

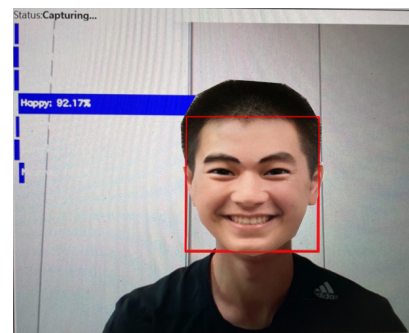


Figure 8. Inspection Analysis Screen

4.2 결과분석

피검자가 질문에 대해 응답할 시 가장 많이 노출하는 감정을 최빈값으로 가정하여 수치 종합본에 기입하였다. 그러나, 질문에 응답 도중 특이 감정이 노출되었을 경우 이를 포착 후 해당 답변을 중심으로 특이 감정의 이유를 확인하는 작업을 같이 진행하였다. <Figure 9>는 예시로 question 1의 Angry의 감정에 대해 분석한 히스토그램과 question 1의 heatmap 시각화자료이다. 히스토그램에서 x축은 감정의 수치, y축은 감정의 수치 빈도를 나타낸다. 히스토그램을 통해 피검자들이 나타내는 각 질문별 감정들의 수치에 대한 빈도를 파악할 수 있

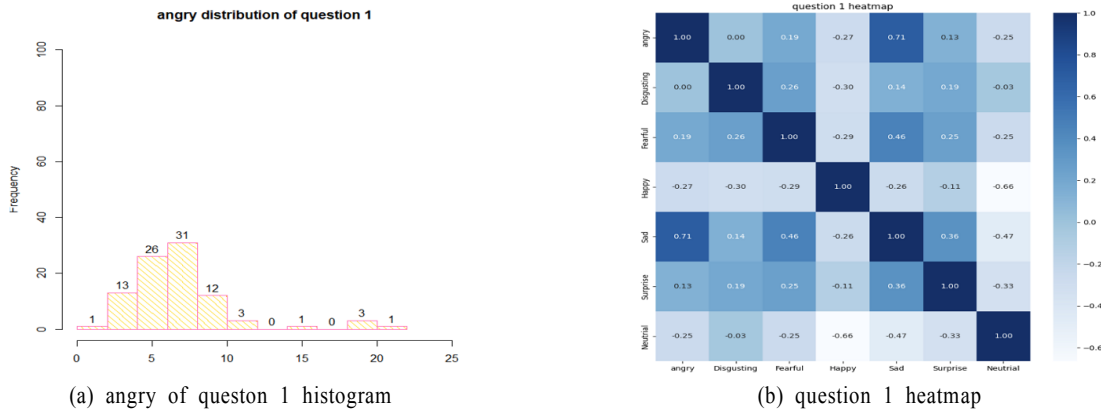


Figure 9. Question 1 Histogram & Heatmap

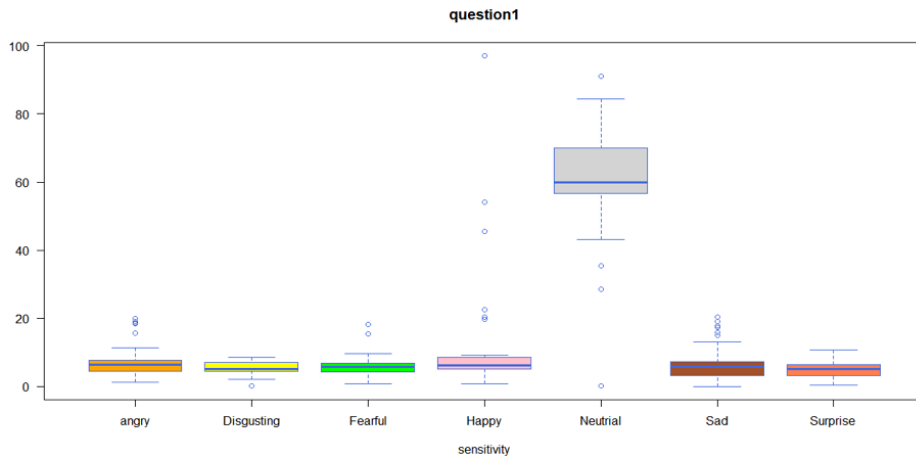


Figure 10. Question 1 Box-plot

다. question 1의 angry는 4~10%의 수치가 피검자들의 평균적인 반응을 나타내는 것을 확인할 수 있다. 또한, heatmap은 각 질문별 감정들의 상관관계를 나타낸다. question 1에서 angry와 sad가 0.71로 가장 강한 상관관계를 나타냄을 확인할 수 있다. 이를 통해 피검자들이 가장 강한 상관관계를 갖는 감정을 드러내는 것과 별개로 다른 상관관계를 갖는 감정들이 나타난다면 관심을 갖고 행동을 관찰해야 하는 간부로 분류하는데 도움이 되는 자료라고 할 수 있다.

또한, <Figure 10>은 question 1의 결과를 Box-plot으로 분석한 결과이다. 해당 자료와 같이 각 질문마다 Box-plot을 통해 나타낸다면 이상치를 쉽게 식별할 수 있다. 이상치로 분류된 인원에 대해서는 앞서 언급했던 바와 같이 특이 감정으로 분류한다. 이후, 검사에 불성실하게 응답하여 노출된 감정인지, 과도하게 감정을 노출시키려다 발생한 것인지 등을 확인한다. question 1에서 Happy의 이상치는 답변을 확인해보니 진정성을 요구하는 답변을 기대하는 질문에서 다른 피검자들과는 달리 웃으면서 답변을 제대로 하지 않고 넘어가려는 불성실한 태도를 보여서 이상치로 분류된 인원이었다. 또한, angry, Fearful, Sad의 이상치는 업무 또는 스트레스에 대한 답변을 할 때 분류된 것으로 확인할 수 있었다. 즉 검사를 통해 이러한 이

상치를 탐색하고 원인을 살펴보는 과정을 진행한다면 본 연구의 목적을 달성할 수 있을 것으로 판단된다.

또한, <Figure 11>은 question 1을 산점도행렬 및 상관관계를 분석한 자료이다. 해당 자료를 통해 앞서 설명했던 히스토그램과 heatmap의 자료를 한 눈에 쉽게 확인할 수 있다.

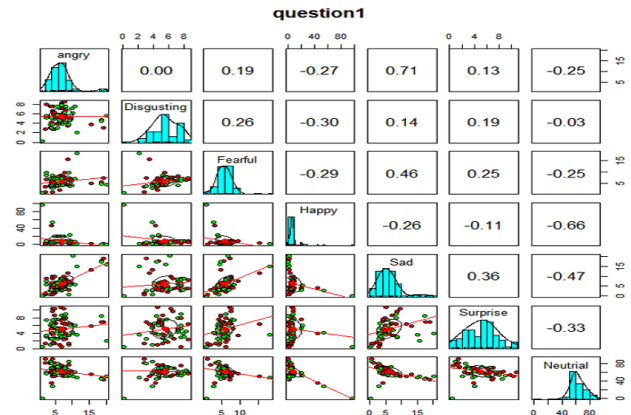


Figure 11. Question 1 Scatter-plot

5. 결 론

본 연구는 신인성검사 2.0에서 사용되고 있는 질문들을 바탕으로 하여 비대면 심리검사방식을 위해 총 10개 질문으로 정리 및 종합 후 피검사 총 40명을 대상으로 검사의 왜곡 가능성 및 성실도 체크를 위해 AI를 활용한 비대면 인터뷰 방식의 검사를 진행하였다. 피검사들을 대상으로 얼굴표정을 통해 감정분석한 결과 히스토그램과 Box plot을 통해 각 질문별 피검사들이 나타내는 감정들을 확인할 수 있었다. 이러한 결과를 기준으로 하여 특이 감정을 나타내는 값을 이상치로 분류하였다. 분류된 인원들은 관심 인원으로 분류하고 해당 인원들에 대해 이상치로 분류된 원인을 확인하였다. 이상치로 분류된 총 2가지 원인으로 정리할 수 있다. 첫 번째는, 검사에 불성실하게 대답하여 검사에서 평균적으로 피검사들이 표현하는 감정 외에 다른 감정을 나타내는 경우이다. 이를 통해 향후 검사에 있어서 성실하게 임했는지에 대한 여부를 확인하는 과정을 정립할 수 있다. 두 번째는, 특정 질문에서 업무 또는 인간관계에 대한 어려움을 호소할 때 순간 나타나는 분노, 슬픔 등의 감정이 이상치로 분류될 수 있었다. 이와 같이 검사 간 피검사들이 나타내는 감정의 수치들을 기준수치로 하여 이보다 더 높거나 적게 나오는 수치들은 오히려 검사결과를 왜곡하기 위해 과도하게 표정 연기를 하는 인원으로 의심하고 주의 깊게 더 관찰을 해야 하는 인원으로 분류할 수 있다. 검사결과를 통해 heatmap 또는 scatter-plot으로 나타내어 각 질문별 드러나는 감정들의 상관관계를 확인하였다. 이를 통해 각 질문별로 나타나는 감정들의 상관관계를 통해 특정 감정이 드러날 때 동반되어 나타나야 하는 감정들을 확인할 수 있는데, 이를 활용하여 결과를 왜곡하려는 인원이 과도하게 연기를 통해 감정을 드러낸다면 오히려 해당 질문에서 상관관계를 가지며 나타나야 할 감정들이 드러나지 않고 다른 감정들이 상관관계를 가지고 드러난다면 해당 인원들을 주의 깊게 관심을 가지고 관찰해야 하는 인원들로 분류하여 관리하는데 정보를 제공할 수 있다. 본 연구는 얼굴표정 감정분석을 통해 더 정확한 검사의 성실도 체크와 결과 왜곡 가능성 식별을 할 수 있다는 가능성을 제시한 점에서 연구의 의의를 도출할 수 있으나 현재 복무 중인 장병 중 소수의 인원을 대상으로 측정된 데이터를 제시한 것으로 인해 향후 검사 인원 확대가 필요하다. 이를 통해 표본을 확대한다면 질문별 기대하는 감정들의 평균치가 더 신뢰성을 갖출 것으로 판단된다. 또한, 얼굴표정 감정분석 이외에 피검자가 답변하는 것을 텍스트화하고 이를 감성 분석하는 텍스트 감성 분석과 우울증이 있는 환자들의 음성을 활용한 주파수 대조 방법을 앙상블하여 멀티-모달 방법을 구축한다면 더 정확하게 본 연구의 목적을 달성할 수 있을 것으로 판단된다. 그러나, 장병들의 검사 성실도 확인과 검사결과 왜곡 가능성 확인을 할 수 있다는 기초 모델과 과학적인 방법을 제시했다는 점에서 본 연구의 의의를 들 수 있다. 현역복무부적합 인원들이 꾸준히 증가하고 있으며, 극단적인 선택에 이르는 장병들이 매년 증가

하고 있는 추세에서 현재 군이 사용하고 있는 신인성검사 2.0 체계 내 복무적응도 검사의 제한사항을 극복함으로써 비전투손실 예방에 적극 기여할 수 있을 것으로 기대된다.

참고문헌

- Ekman, P. (2014), *Facial expression reading skills*, Cheonglim Publishing, Seoul, Korea.
- Hong, S. K. (2006), A Study on Pattern of Facial Expression Presentation in Character Animation, *The Korea Contents Association*, 6(8), 165-174.
- Jung, B. H. (2017), A Study on the Application of Literature to Military Unit New Personality Test: A Study on the Fingerprint and Psychological Characteristics, a master's thesis, Daejin University Graduate School of Unification.
- Kim, I. G. (2018), *A study on the Adaptation of barracks in 2018*, KIDA, Seoul, Korea.
- Kim, S. G. (2021), A Study on the Correlation of New personality Test 2.0 and Biometric Indicators: Focused on the Service Adaptability Test 2.0 Results, a master's thesis, Seoul National University of Science and Technology.
- Kwon, S. M. (2023), 80% of all soldiers who are not eligible for active duty are "mental illness and maladjustment", *Daily economy*, 2023.08.23, <https://www.mk.co.kr/news/society/10815538>.
- LaFrance, M. (2012), *The psychology of laughter*, Jbooks Publishing, Seoul, Korea.
- Liu, Z., Mao, H., Wu, C., Feichtenhofer, C., Darrell, T., and Xie, S. (2020), *A ConvNet for the 2020s*, *Facebook AI Research (FAIR)*, UC Berkeley.
- Nam, J. H. (2019), Design of Ontology Implementation Process for the Identification of Non-serviceable Soldiers, *Korean Society of Computer Information*, 27(2), 171-174.
- Park, E. H. (2022), Comparison of improvement in items for military serviceadaptability test and analysis between militaryaptitude-adjustment prediction techniques, *Proceedings of KIIT Conference*, 122-123.
- Snowden, R. (2013), *The foundation of visual psychology*, Hakjisa Publishing, Seoul, Korea.
- Woo, H. S. (1992), *How to train a Keras model on TFRecord files*, Keras running day, Korea University

저자소개

조성진: 육군3사관학교 무기시스템학과에서 2016년 학사, 국방대학교에서 국방과학학과 군사운영분석전공으로 석사과정을 취득하고 2024년부터 육군3사관학교 국방시스템학과 교수로 재직하고 있다. 연구분야는 최적화, 딥러닝, 데이터마이닝이다.

김수환: 육군사관학교에서 1996년 학사, 한국과학기술원 산업공학과에서 2004년 석사학위를 취득하고 Texas A&M 대학교에서 2013년에 산업공학 박사학위를 취득하였다. 현재 국방대학교 국방과학학부 교수로 재직하고 있다. 연구분야는 조합최적화, 인공지능, 비용추정이다.