



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0015183
(43) 공개일자 2025년02월03일

(51) 국제특허분류(Int. Cl.)
G06V 10/80 (2022.01) G06V 10/75 (2022.01)
G06V 10/764 (2022.01)
(52) CPC특허분류
G06V 10/806 (2023.08)
G06F 18/24133 (2023.01)
(21) 출원번호 10-2023-0096268
(22) 출원일자 2023년07월24일
심사청구일자 2023년07월24일

(71) 출원인
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
(72) 발명자
이상철
인천광역시 연수구 아트센터대로97번길 76, 1402동 805호(송도동, 송도 더샵 하버뷰 14단지)
박서형
서울특별시 노원구 동일로245길 162, 101동(상계동, 은빛1단지아파트)
(74) 대리인
양성보

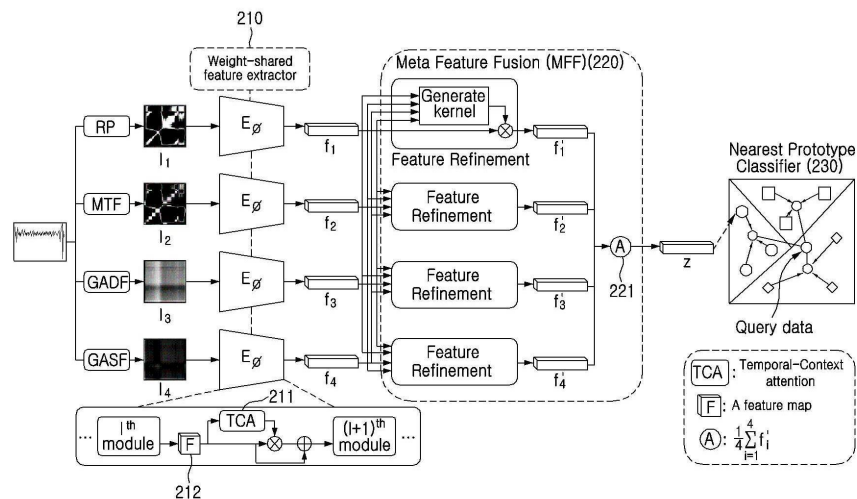
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 시계열 데이터 분류를 위한 데이터 융합 방법 및 시스템

(57) 요약

시계열 데이터 분류를 위한 데이터 융합 방법 및 시스템이 제시된다. 본 발명에서 제안하는 시계열 데이터 분류를 위한 데이터 융합 방법은 특징 추출부를 통해 입력 이미지에 대해 TCA(Temporal-Context Attention)를 이용하여 특징 맵의 글로벌 컨텍스트를 통합하고 다른 픽셀과 정보 의존성을 갖는 해당 픽셀을 강조하여 특징을 추출하는 단계, 특징 융합부를 통해 상기 추출된 특징에 대해 MFF(Meta Feature Fusion)를 이용하여 교차 모달리티 특징을 기반으로 커널을 생성하고, 생성된 커널을 이용하여 각 특징을 세분화하고 정제된 특징을 융합하는 단계 및 분류부를 통해 상기 융합된 특징에 대하여 임베딩 공간에서 가장 가까운 클래스 프로토타입으로 분류하는 단계를 포함한다.

대표도



(52) CPC특허분류

G06V 10/75 (2023.08)

G06V 10/764 (2023.08)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711192607
과제번호	2021R1A2C2010893
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	중견연구
연구과제명	[Ezbaro] 희소 영상에 특화된 영상 분해-합성을 통한멀티도메인 기반 관심 지표 및
객체 검출 기술 연구	
기 여 율	1/2
과제수행기관명	인하대학교
연구기간	2023.03.01 ~ 2024.02.29

이 발명을 지원한 국가연구개발사업

과제고유번호	1711179347
과제번호	00155915
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신인재양성(교육운영수익)
연구과제명	[Ezbaro][정부+민간] 인공지능융합혁신인재양성(2차년도)
기 여 율	1/2
과제수행기관명	인하대학교 산학협력단
연구기간	2023.01.01 ~ 2023.12.31

명세서

청구범위

청구항 1

특징 추출부를 통해 입력 이미지에 대해 TCA(Temporal-Context Attention)를 이용하여 특징 맵의 글로벌 컨텍스트를 통합하고 다른 픽셀과 정보 의존성을 갖는 해당 픽셀을 강조하여 특징을 추출하는 단계;

특징 융합부를 통해 상기 추출된 특징에 대해 MFF(Meta Feature Fusion)를 이용하여 교차 모달리티 특징을 기반으로 커널을 생성하고, 생성된 커널을 이용하여 각 특징을 세분화하고 정제된 특징을 융합하는 단계; 및

분류부를 통해 상기 융합된 특징에 대하여 임베딩 공간에서 가장 가까운 클래스 프로토타입으로 분류하는 단계를 포함하는 데이터 융합 방법.

청구항 2

제1항에 있어서,

상기 특징 추출부를 통해 입력 이미지에 대해 TCA를 이용하여 특징 맵의 글로벌 컨텍스트를 통합하고 다른 픽셀과 정보 의존성을 갖는 해당 픽셀을 강조하여 특징을 추출하는 단계는,

상기 입력 이미지를 인코딩하기 위해 TCA를 이용하여 입력 이미지의 픽셀 간 정보 의존성을 캡처하기 위한 bi-GRU(bi-directional gated recurrent unit)를 통해 글로벌 컨텍스트를 통합하고, FSL(Few-shot learning)을 통해 시계열의 인코딩된 이미지를 훈련시키는

데이터 융합 방법.

청구항 3

제1항에 있어서,

상기 특징 융합부를 통해 상기 추출된 특징에 대해 MFF(Meta Feature Fusion)를 이용하여 교차 모달리티 특징을 기반으로 커널을 생성하고, 생성된 커널을 이용하여 각 특징을 세분화하고 정제된 특징을 융합하는 단계는,

상기 추출된 서로 다른 복수의 특징을 쌍으로 곱하여 교차 모달리티 특징을 획득하고, 상기 교차 모달리티 특징을 평균화한 후 상기 교차 모달리티 특징과 적응형 커널을 이용하여 각 특징을 차별적으로 정제하며, 상기 정제된 각 특징을 평균화하여 최종 특징을 출력하는

데이터 융합 방법.

청구항 4

제1항에 있어서,

상기 분류부를 통해 상기 융합된 특징에 대하여 임베딩 공간에서 가장 가까운 클래스 프로토타입으로 분류하는 단계는,

상기 특징 추출부 및 상기 특징 융합부를 통해 특징 벡터를 계산하고, 미리 정해진 임베딩 함수를 이용하여 클래스 특징을 획득하여, NPC(Nearest Prototype Classifier)를 통해 클래스에 대한 클래스 프로토타입을 계산하고,

임베딩 공간에서 쿼리 샘플의 특징을 나타내는 각 클래스에 대한 레이블 분포를 생성하며, 클래스 특징 및 클래스 프로토타입 간의 유클리드거리에 따라 해당 클래스 특징을 가까운 클래스 프로토타입으로 분류하는

데이터 융합 방법.

청구항 5

입력 이미지에 대해 TCA(Temporal-Context Attention)를 이용하여 특징 맵의 글로벌 컨텍스트를 통합하고 다른

픽셀과 정보 의존성을 갖는 해당 픽셀을 강조하여 특징을 추출하는 특징 추출부;

상기 추출된 특징에 대해 MFF(Meta Feature Fusion)를 이용하여 교차 모달리티 특징을 기반으로 커널을 생성하고, 생성된 커널을 이용하여 각 특징을 세분화하고 정제된 특징을 융합하는 특징 융합부; 및

상기 융합된 특징에 대하여 임베딩 공간에서 가장 가까운 클래스 프로토타입으로 분류하는 분류부를 포함하는 데이터 융합 시스템.

청구항 6

제5항에 있어서,

상기 특징 추출부는,

상기 입력 이미지를 인코딩하기 위해 TCA를 이용하여 입력 이미지의 픽셀 간 정보 의존성을 캡처하기 위한 bi-GRU(bi-directional gated recurrent unit)를 통해 글로벌 컨텍스트를 통합하고, FSL(Few-shot learning)을 통해 시계열의 인코딩된 이미지를 훈련시키는

데이터 융합 시스템.

청구항 7

제5항에 있어서,

상기 특징 융합부는,

상기 추출된 서로 다른 복수의 특징을 쌍으로 곱하여 교차 모달리티 특징을 획득하고, 상기 교차 모달리티 특징을 평균화한 후 상기 교차 모달리티 특징과 적응형 커널을 이용하여 각 특징을 차별적으로 정제하며, 상기 정제된 각 특징을 평균화하여 최종 특징을 출력하는

데이터 융합 시스템.

청구항 8

제5항에 있어서,

상기 분류부는,

상기 특징 추출부 및 상기 특징 융합부를 통해 특징 벡터를 계산하고, 미리 정해진 임베딩 함수를 이용하여 클래스 특징을 획득하여, NPC(Nearest Prototype Classifier)을 통해 클래스에 대한 클래스 프로토타입을 계산하고,

임베딩 공간에서 쿼리 샘플의 특징을 나타내는 각 클래스에 대한 레이블 분포를 생성하며, 클래스 특징 및 클래스 프로토타입 간의 유클리드거리에 따라 해당 클래스 특징을 가까운 클래스 프로토타입으로 분류하는

데이터 융합 시스템.

발명의 설명

기술 분야

[0001] 본 발명은 시계열 데이터 분류를 위한 데이터 융합 방법 및 시스템에 관한 것이다.

배경 기술

[0002] 시계열은 재무 기록, 심리적 신호 및 날씨 측정과 같은 시간 데이터 객체의 다양한 분야를 나타내는 시간 영역의 데이터 포인트 시퀀스이다. 현재 디지털 시대에 시계열 데이터 분석은 예측, 환경 모니터링 및 결함 진단을 포함하여 도메인 전반에 걸쳐 널리 사용되어 왔다. 지난 수십 년 동안 시계열 분류(Time-Series Classification; TSC)는 시계열 데이터 분석의 필수 작업이 되었으며, 시계열 데이터를 정확하게 분류하기 위한 다양한 방법이 제안되었다.

[0003] 초기 알고리즘은 거리 기반, 사전 기반, 간격 기반 및 shapelet 기반과 같은 시계열 데이터를 표현하는 데 중점

을 둔다. 여러 분류기를 결합한 앙상블 기반 방법은 최근 최첨단(State-Of-The-Art; SOTA) 성능을 보여주었다. 그러나 이러한 방법은 엄청난 계산 복잡성과 수작업 특징 엔지니어링으로 인해 실제 문제에 적용하기에는 비현실적이다. 컴퓨터 비전 작업에서 딥 러닝의 성공으로, 딥 러닝은 TSC 커뮤니티에서 종단 간 TSC를 위한 정확한 특징을 자동으로 추출하는 데 인기를 얻었다. LSTMFCN은 장단기 메모리(Long Short-Term Memory; LSTM) 네트워크와 컨볼루션 신경망(Convolutional Neural Networks; CNN)을 사용하여 시계열에서 특징을 추출한다. 인셉션 타임(InceptionTime)은 5개의 인셉션(Inception) 기반 CNN의 앙상블을 사용한다.

[0004] 그럼 불구하고, 딥 러닝 기반 접근 방식은 대규모 데이터에 크게 의존하며, 이는 많은 양의 레이블이 지정된 데이터를 사용할 수 없을 때 과적합으로 이어질 수 있다. 대규모 데이터를 수집하려면 의료 및 예지 컴퓨팅 도메인과 같은 여러 도메인에서 충분한 데이터 소스와 높은 주석 비용이 필요하다. 따라서, 이 문제를 해결하기 위해 데이터 확대, 준지도 학습 및 전이 학습을 포함한 여러 접근 방식이 제안되었다. 그러나, 그것들은 여전히 극도로 제한된 수의 훈련 샘플, 예를 들어 클래스당 한 개 또는 다섯 개의 샘플로 새로운 작업에 적응하는 데 어려움을 겪고 있다. 이러한 문제(즉, 퓨샷 시나리오)는 컴퓨터 비전과 자연어 처리(Natural Language Processing; NLP)에서 이미 논의되었다.

[0005] 퓨샷 학습(Few-Shot Learning; FSL)은 레이블이 있는 훈련 샘플 몇 개만으로 일반화하는 방법을 학습하여 이 문제를 해결하는 데 사용된다. 일반적으로 FSL에는 두 가지 주요 흐름이 있다. 다양한 새로운 작업에 신속하게 최적화하는 것을 대상으로 하는 메타 학습 기반 접근 방식과 대상 태스크의 표현을 학습하고 거리 함수를 사용하여 데이터를 분류하는 지표 기반 접근 방식이다. 좋은 표현을 학습하는 것은 두 접근 방식 모두 차별적 특징을 사용하여 정확하게 분류하는 데 중요하다.

[0006] 1D 시계열을 2D 영상으로 인코딩하는 것은 주기성 및 시간 의존성과 같이 원래 데이터에서 관찰되지 않는 숨겨진 특징을 드러내 시계열의 표현 능력을 높이는 방법 중 하나이다. 대표적인 이미지 인코딩 방법 중 세 가지 방법이 TSC에서 주목할 만한 성능을 보여주었다: 세 가지 방법은 반복 플롯(Recurrence Plot; RP), 마르코프 전이 필드(Markov Transition Field; MTF), 그리고 Gramian 각도 합계/차이 필드(GASF(Gramian Angular Summation Field) 및 GADF(Gramian Angular Difference Field)이다.

[0007] 도 1은 종래기술에 따른 RP, MTF, GASF 및 GADF를 사용하여 시계열을 이미지로 인코딩하는 예시를 설명하기 위한 도면이다.

[0008] 각 데이터는 UCR(University of California, Riverside) 아카이브의 세 가지 데이터 세트(도 1(a): Crop, 도 1(b): Phoneme, (c): PigCVP)에서 가져온다.

[0009] 도 1은 해당 방법을 시계열에 적용하는 예시를 보여준다. 그럼에도 불구하고, 각각은 MTF의 역 맵의 불확실성과 RP의 경향 혼란 문제와 같은 자체적인 한계를 가지고 있다. 이러한 문제를 해결하기 위해, 종래기술에서는 복수의 이미지(즉, 멀티모달 융합)를 융합하면 단일 이미지를 사용하는 것보다 더 높은 성능을 달성할 수 있다는 것을 입증했다. 그러나 멀티모달 특징을 효과적으로 융합하는 방법은 여전히 어렵다.

선행기술문헌

특허문헌

[0010] (특허문헌 0001) [1] F. Visin, K. Kastner, K. Cho, M. Matteucci, A. Courville, and Y. Bengio, "ReNet: A recurrent neural network based alternative to convolutional networks," 2015, arXiv:1505.00393.

발명의 내용

해결하려는 과제

[0011] 본 발명이 이루고자 하는 기술적 과제는 시계열을 다른 유형의 이미지(다시 말해, 반복 플롯, 마르코프 전이 필드, Gramian 각도 합계/차등 필드)로 인코딩하고 FSL 절차를 사용하여 이러한 이미지를 모델로 훈련시키는 퓨샷 TSC를 위한 시계열 데이터 분류를 위한 데이터 융합 방법 및 시스템을 제공하는데 있다.

[0012] 본 발명의 실시예에 따르면, 이미지의 표현 능력을 최대화하기 위해 시간적 컨텍스트 주의(Temporal-Context Attention; TCA) 및 메타 특징 융합(Meta-Feature Fusion; MFF)을 제안한다.

과제의 해결 수단

- [0013] 일 측면에 있어서, 본 발명에서 제안하는 시계열 데이터 분류를 위한 데이터 융합 방법은 특징 추출부를 통해 입력 이미지에 대해 TCA(Temporal-Context Attention)를 이용하여 특징 맵의 글로벌 컨텍스트를 통합하고 다른 픽셀과 정보 의존성을 갖는 해당 픽셀을 강조하여 특징을 추출하는 단계, 특징 융합부를 통해 상기 추출된 특징에 대해 MFF(Meta Feature Fusion)를 이용하여 교차 모달리티 특징을 기반으로 커널을 생성하고, 생성된 커널을 이용하여 각 특징을 세분화하고 정제된 특징을 융합하는 단계 및 분류부를 통해 상기 융합된 특징에 대하여 임베딩 공간에서 가장 가까운 클래스 프로토타입으로 분류하는 단계를 포함한다.
- [0014] 상기 특징 추출부를 통해 입력 이미지에 대해 TCA를 이용하여 특징 맵의 글로벌 컨텍스트를 통합하고 다른 픽셀과 정보 의존성을 갖는 해당 픽셀을 강조하여 특징을 추출하는 단계는 상기 입력 이미지를 인코딩하기 위해 TCA를 이용하여 입력 이미지의 픽셀 간 정보 의존성을 캡처하기 위한 bi-GRU(bi-directional gated recurrent unit)를 통해 글로벌 컨텍스트를 통합하고, FSL(Few-shot learning)을 통해 시계열의 인코딩된 이미지를 훈련시킨다.
- [0015] 상기 특징 융합부를 통해 상기 추출된 특징에 대해 MFF(Meta Feature Fusion)를 이용하여 교차 모달리티 특징을 기반으로 커널을 생성하고, 생성된 커널을 이용하여 각 특징을 세분화하고 정제된 특징을 융합하는 단계는 상기 추출된 서로 다른 복수의 특징을 쌍으로 곱하여 교차 모달리티 특징을 획득하고, 상기 교차 모달리티 특징을 평균화한 후 상기 교차 모달리티 특징과 적응형 커널을 이용하여 각 특징을 차별적으로 정제하며, 상기 정제된 각 특징을 평균화하여 최종 특징을 출력한다.
- [0016] 상기 분류부를 통해 상기 융합된 특징에 대하여 임베딩 공간에서 가장 가까운 클래스 프로토타입으로 분류하는 단계는 상기 특징 추출부 및 상기 특징 융합부를 통해 특징 벡터를 계산하고, 미리 정해진 임베딩 함수를 이용하여 클래스 특징을 획득하여, NPC(Nearest Prototype Classifier)를 통해 클래스에 대한 클래스 프로토타입을 계산하고, 임베딩 공간에서 쿼리 샘플의 특징을 나타내는 각 클래스에 대한 레이블 분포를 생성하며, 클래스 특징 및 클래스 프로토타입 간의 유클리드거리에 따라 해당 클래스 특징을 가까운 클래스 프로토타입으로 분류한다.
- [0017] 또 다른 일 측면에 있어서, 본 발명에서 제안하는 시계열 데이터 분류를 위한 데이터 융합 시스템은 입력 이미지에 대해 TCA(Temporal-Context Attention)를 이용하여 특징 맵의 글로벌 컨텍스트를 통합하고 다른 픽셀과 정보 의존성을 갖는 해당 픽셀을 강조하여 특징을 추출하는 특징 추출부, 상기 추출된 특징에 대해 MFF(Meta Feature Fusion)를 이용하여 교차 모달리티 특징을 기반으로 커널을 생성하고, 생성된 커널을 이용하여 각 특징을 세분화하고 정제된 특징을 융합하는 특징 융합부 및 상기 융합된 특징에 대하여 임베딩 공간에서 가장 가까운 클래스 프로토타입으로 분류하는 분류부를 포함한다.

발명의 효과

- [0018] 본 발명의 실시예들에 따르면 시계열 데이터 분류를 위한 데이터 융합 방법 및 시스템을 통해 시계열을 다른 유형의 이미지(다시 말해, 반복 플롯, 마르코프 전이 필드, Gramian 각도 합계/차등 필드)로 인코딩하고 FSL 절차를 사용하여 이러한 이미지를 모델로 훈련시킬 수 있다. 또한, 시간적 컨텍스트 주의(Temporal-Context Attention; TCA) 및 메타 특징 융합(Meta-Feature Fusion; MFF)을 통해 이미지의 표현 능력을 최대화할 수 있다. 본 발명의 실시예에 따르면, 각 이미지의 다양한 기능을 통해 모델은 풍부한 정보를 학습할 수 있다. 또한, 시간적 컨텍스트 주의(Temporal-Context Attention; TCA) 및 메타 특징 융합(Meta-Feature Fusion; MFF)을 통해 이러한 이미지의 표현 능력을 최대화할 수 있다. 본 발명의 실시예들에 따르면, TCA는 특징 맵의 글로벌 컨텍스트를 통합하고 다른 픽셀과 정보 관련성이 있는 픽셀을 강조한다. 특징을 추출한 후, MFF는 교차 모달리티 특징을 기반으로 생성된 서로 다른 커널을 사용하여 각 특징을 세분화하고 정제된 특징을 융합할 수 있다.

도면의 간단한 설명

- [0019] 도 1은 종래기술에 따른 RP, MTF, GASF 및 GADF를 사용하여 시계열을 이미지로 인코딩하는 예시를 설명하기 위한 도면이다.
- 도 2는 본 발명의 일 실시예에 따른 시계열 데이터 분류를 위한 데이터 융합 시스템의 구성을 나타내는 도면이다.
- 도 3은 본 발명의 일 실시예에 따른 시간적 컨텍스트 주의를 설명하기 위한 도면이다.

도 4는 본 발명의 일 실시예에 따른 메타 특징 융합의 전체 기능 개선 프로세스를 설명하기 위한 도면이다.

도 5는 본 발명의 일 실시예에 따른 시계열 데이터 분류를 위한 데이터 융합 방법을 설명하기 위한 흐름도이다.

도 6은 본 발명의 일 실시예에 따른 메타 특징 융합 알고리즘을 나타내는 도면이다.

도 7은 본 발명의 일 실시예에 따른 FSL에서 네트워크를 훈련시키는 알고리즘을 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0020] 딥 러닝은 종단 간 시계열 분류(Time-Series Classification; TSC)에 널리 채택되었다. 그러나 딥 러닝의 효과는 대규모 데이터에 크게 의존한다. 따라서 딥 러닝은 레이블이 지정된 샘플이 거의 없을 때 과적합되기 쉽다. 퓨샷 학습(Few-Shot Learning; FSL)은 훈련 샘플이 거의 없는 새로운 작업(예를 들어, 클래스당 샘플 1개 또는 5개)으로 일반화하는 방법을 학습하여 이 문제를 해결하는 것을 목표로 한다. FSL은 차별적 특징을 사용하여 정확하게 분류하기 위해 좋은 표현을 학습하는 것이 중요하다.
- [0021] 본 발명에서, 시계열을 다른 유형의 이미지(다시 말해, 반복 플롯, 마르코프 전이 필드, Gramian 각도 합계/차이 필드)로 인코딩하고 FSL 절차를 사용하여 이러한 이미지를 모델로 훈련시키는 퓨샷 TSC를 위한 프레임워크를 제안한다. 각 이미지의 다양한 기능을 통해 모델은 풍부한 정보를 학습할 수 있다. 또한 이러한 이미지의 표현 능력을 최대화하기 위해 시간적 컨텍스트 주의(Temporal-Context Attention; TCA) 및 메타 특징 융합(Meta-Feature Fusion; MFF)을 제안한다. TCA는 특징 맵의 글로벌 컨텍스트를 통합하고 다른 픽셀과 정보 관련성이 있는 픽셀을 강조한다. 특징을 추출한 후, MFF는 교차 모달리티 특징을 기반으로 생성된 서로 다른 커널을 사용하여 각 특징을 세분화하고 정제된 특징을 융합한다. 이하, 본 발명의 실시 예를 첨부된 도면을 참조하여 상세하게 설명한다.
- [0023] 본 발명에서는 소수의 훈련 샘플로 학습하기 위해 퓨샷 시나리오(즉, N-way K-shot 문제)를 가정하여 시계열을 이미지로 인코딩하고 이러한 이미지를 수많은 TSC 태스크에서 네트워크로 훈련시키는 퓨샷 TSC를 위한 새로운 프레임워크를 제안한다.
- [0024] 더욱 상세하게는, 본 발명의 실시예에 따른 프레임워크는 먼저 대상 태스크에 대한 풍부한 정보를 얻기 위해 여러 이미지(RP, MTF, GASF 및 GADF)에 시계열을 인코딩한다. 또한, 본 발명의 실시예에 따른 프레임워크에는 이미지의 표현 능력을 최대화하기 위한 두 가지 주요 구성 요소, 즉 시간적 컨텍스트 주의(Temporal-Context Attention; TCA) 및 메타 특징 융합(Meta-Feature Fusion; MFF)이 있다. 특징 추출부에서 TCA는 양방향 게이트 반복 유닛(bi-GRU)를 사용하여 특징 맵의 글로벌 컨텍스트를 통합하고 다른 픽셀과 정보 관련성이 있는 픽셀을 강조한다. 특징 융합부는 서로 다른 커널을 사용하여 각 이미지의 출력 기능을 정제한 다음 정제된 특징을 융합한다. 각 이미지의 커널은 교차 모달리티 특징을 사용하여 생성되고 기능을 보다 차별적으로 개선하도록 훈련된다. 마지막으로, FSL에서 지표 기반 접근 방식으로 널리 사용되는 가장 가까운 프로토타입 분류기(Nearest Prototype Classifier; NPC)는 임베딩 공간에서 가장 가까운 클래스 프로토타입에 따라 테스트 샘플을 분류한다.
- [0025] TSC는 시계열을 정확하게 분류하기 위해 차별적 표현을 학습하는 것을 목표로 하며, 이 목표를 달성하기 위해 다양한 방법이 제안되었다. 근접 포레스트와 같은 거리 기반 알고리즘은 두 계열 사이의 탄성 거리를 측정한다. 반면에 BOSS 및 WEISEL와 같은 사전 기반 접근 방식은 시계열에서 패턴의 발생 빈도에 기초한 표현을 사용한다. 또한 TSF 및 RISE와 같은 간격 기반 알고리즘은 관련 없는 관찰에 의해 가려질 수 있는 시간적 특징을 결정하기 위해 시계열 간격에 대한 특징을 도출한다. 또한 STC와 같은 shapelet 기반 접근 방식은 차별적인 서브 계열, 이른바 'shapelets'를 결정하려고 시도한다. 최근 앙상블 기반 방법은 HIVECOTE, TS-Chief, Rocket와 같은 여러 분류기를 결합하여 SOTA 성능을 보여주었다. 그러나 이러한 방법은 여러 분류기와 수작업 특징 엔지니어링을 사용하기 때문에 계산 집약적이다.
- [0026] CNN은 이미지 인식 및 NLP와 같은 다양한 분야에 혁명을 일으켰다. 마찬가지로, 점점 더 많은 수의 작업이 TSC에 CNN을 적용했다. 다중 스케일 CNN은 UCR 아카이브에서 검증된 최초의 아키텍처 중 하나이다. 완전 컨볼루션 신경망(Fully Convolutional Neural Networks; FCN)은 계층 풀링 없이 개선을 달성한다. 최근 ResNet은 잔차 연결 ResNet을 사용하여 성능을 더욱 향상시킨다. LSTM-FCN은 LSTM과 CNN을 결합하여 시계열 특징을 추출한다. 인셉션 타임(InceptionTime)은 5개의 인셉션 기반 CNN의 앙상블을 사용한다. 최근에는 OS-CNN이 OS 블록을 사용하여 1D-CNN에 대해 다양한 규모의 수신 필드를 자동으로 채택한다. 그럼에도 불구하고, 딥 러닝 접근법의 효과

는 대규모 데이터에 크게 의존하며, 이는 훈련 샘플이 거의 없을 때 과적합으로 이어진다.

- [0027] 시계열 분류에서 희소 데이터 문제 처리에 있어서, 부족한 데이터 문제를 해결하기 위해 몇 가지 TSC 접근 방식이 제안되었다. 데이터 확대는 자르기, 슬라이싱 및 뒤틀림과 같은 조작된 데이터를 생성하거나 심층 생성 모델을 사용하여 거의 현실적인 데이터를 만든다. 일부 종래기술은 예측 또는 레이블이 없는 데이터 간의 시간적 관계 예측과 같은 보조 작업으로 네트워크를 훈련시킴으로써 준지도 학습을 활용한다. 또한, 몇몇 연구는 사전 훈련된 네트워크를 사용하여 소스 도메인에서 지식을 이동함으로써 전이 학습을 채택한다. 그러나 이러한 방법은 훈련 샘플이 거의 없이 새로운 작업이 제공될 때 성능을 보장하는 데 어려움이 있다.
- [0028] 최근 FSL은 이미 컴퓨터 비전과 NLP에서 이러한 문제를 해결하기 위해 사용되었다. 최근 TSC에서 동일한 문제를 해결하기 위해 여러 연구에서 FSL을 채택했다. DPSN은 SFA 특징 추출을 채택하고 가장 가까운 프로토타입 네트워크를 통해 데이터를 분류한다. Narwariya는 메타 학습 절차를 사용하여 다양한 몇 번의 TSC 작업을 해결하도록 CNN을 훈련했다. 본 발명의 실시예에 따른 프레임워크는 퓨샷 TSC 시나리오에 대한 높은 일반화 능력을 달성하는 것을 목표로 한다.
- [0029] 본 발명의 실시예에 따르면, FSL 절차를 사용하여 시계열의 인코딩된 이미지를 네트워크에 훈련시킴으로써 훈련 샘플이 거의 없는 새로운 태스크에 빠르게 적응하는 퓨샷 TSC를 위한 프레임워크를 제안한다.
- [0030] 본 발명의 실시예에 따르면, 인코딩된 이미지의 숨겨진 특징을 효과적으로 학습하기 위해 TCA와 MFF를 제안한다. TCA는 픽셀 간의 정보 의존성을 캡처하기 위해 bi-GRU를 사용하여 글로벌 컨텍스트를 통합한다. MFF는 교차 모달리티 특징을 사용하여 적응적으로 생성된 커널을 사용하여 특징을 더욱 차별화한다.
- [0031] 본 발명의 실시예에 따르면, 다양한 N-way K-shot 문제에 대한 광범위한 실험은 우리의 프레임워크가 0.34~29.4%의 마진으로 UCR 아카이브의 28개 데이터 세트에서 SOTA 모델을 능가함을 보여준다.
- [0033] 도 2는 본 발명의 일 실시예에 따른 시계열 데이터 분류를 위한 데이터 융합 시스템의 구성을 나타내는 도면이다.
- [0034] 본 발명의 실시예에 따른 시계열 데이터 분류를 위한 데이터 융합 시스템은 특징 추출부(210), 특징 융합부(220) 및 분류부(230)를 포함한다.
- [0035] 본 발명의 실시예에 따른 특징 추출부(210)는 입력 이미지에 대해 TCA(Temporal-Context Attention)(211)를 이용하여 특징 맵(212)의 글로벌 컨텍스트를 통합하고 다른 픽셀과 정보 의존성을 갖는 해당 픽셀을 강조하여 특징을 추출한다.
- [0036] 본 발명의 실시예에 따른 특징 추출부(210)는 상기 입력 이미지를 인코딩하기 위해 TCA를 이용하여 입력 이미지의 픽셀 간 정보 의존성을 캡처하기 위한 bi-GRU(bi-directional gated recurrent unit)를 통해 글로벌 컨텍스트를 통합하고, FSL(Few-shot learning)을 통해 시계열의 인코딩된 이미지를 훈련시킨다.
- [0037] 본 발명의 실시예에 따른 특징 융합부(220)는 상기 추출된 특징에 대해 MFF(Meta Feature Fusion)를 이용하여 교차 모달리티 특징을 기반으로 커널을 생성하고, 생성된 커널을 이용하여 각 특징을 세분화하고 정제된 특징을 융합한다.
- [0038] 본 발명의 실시예에 따른 특징 융합부(220)는 상기 추출된 서로 다른 복수의 특징을 쌍으로 곱하여 교차 모달리티 특징을 획득하고, 상기 교차 모달리티 특징을 평균화한 후 상기 교차 모달리티 특징과 적응형 커널을 이용하여 각 특징을 차별적으로 정제하며, 상기 정제된 각 특징을 평균화하여 최종 특징을 출력한다.
- [0039] 본 발명의 실시예에 따른 분류부(230)는 상기 융합된 특징에 대하여 임베딩 공간에서 가장 가까운 클래스 프로토타입으로 분류한다.
- [0040] 본 발명의 실시예에 따른 분류부(230)는 상기 특징 추출부 및 상기 특징 융합부를 통해 특징 벡터를 계산하고, 미리 정해진 임베딩 함수를 이용하여 클래스 특징을 획득하여, NPC(Nearest Prototype Classifier)를 통해 클래스에 대한 클래스 프로토타입을 계산한다.
- [0041] 본 발명의 실시예에 따른 분류부(230)는 임베딩 공간에서 쿼리 샘플의 특징을 나타내는 각 클래스에 대한 레이블 분포를 생성하며, 클래스 특징 및 클래스 프로토타입 간의 유클리드거리에 따라 해당 클래스 특징을 가까운 클래스 프로토타입으로 분류한다.

[0042] 본 발명의 실시예에 따르면, 클래스 수를 클래스당 N 및 K 샘플로 제한하는 분류 시나리오, 즉 N -way K -shot 문제를 수행한다. 레이블이 있는 큰 훈련 집합 $\mathcal{D}^{train}$ 이 주어지면 예상되는 퓨샷 시나리오를 가정하여 $\mathcal{D}^r \subseteq \mathcal{D}^{train}$ 을 많은 태스크로 무작위로 샘플링한다. 각 $\mathcal{D}^r$ 은 N 개의 클래스, 클래스당 K 개의 레이블이 있는 시계열 및 레이블이 없는 N_Q 개의 시계열로 구성된다. 이와 같이, $N \times K$ 레이블이 있는 데이터는 각각 네트워크를 훈련시키는 지원 집합 $\mathcal{S}^r = (X_i, Y_i)_{i=1}^{N \times K}$, 레이블이 없는 쿼리 집합 $\mathcal{Q}^r = (X_i^*, Y_i^*)_{i=1}^{N_Q}$ 으로 사용된다. 여기서 N_Q 는 $\mathcal{Q}^r$ 에 대해 레이블이 없는 데이터의 수이다. $\mathcal{S}^r$ 와 $\mathcal{Q}^r$ 는 분리되었다. 즉, $\mathcal{S}^r \cap \mathcal{Q}^r = \emptyset$ 이다. 또한, $X_i, X_i^* \in \mathbb{R}^t$ 는 길이 t 의 일변량 시계열을 나타내며 Y_i, Y_i^* 은 지원 및 쿼리 레이블을 나타낸다. 여기서, $Y_i, Y_i^* \in \{1, \dots, N\}$ 이다. 목표는 $\mathcal{Q}^r$ 의 레이블을 정확하게 예측하는 $\mathcal{S}^r$ 에 대한 분류기를 학습하는 것이다. 도 3 및 도 4를 참조하여, TCA 및 MTF에 대하여 더욱 상세히 설명한다.

[0043] 제안하는 방법을 설명하기 전에 먼저 본 발명의 실시예에서 사용되는 이미지 인코딩 방법을 설명한다. 각 이미지의 너비와 높이는 시계열 길이와 동일하다. 이미지를 그레이 스케일로 시각화하여 입력 데이터의 차원과 RGB 이미지에 대한 복잡성을 줄이면서 정확성을 보장한다. $X = \{x_1, x_2, \dots, x_t\} \in \mathbb{R}^t$ 가 설명을 위한 일변량 시계열이라 하자.

[0044] RP는 주기성, 불규칙 순환, 안정성과 같은 시계열의 다양한 시간적 특징을 시각화한다. 핵심 아이디어는 위상 공간에서 어떤 위상 공간 궤적이 거의 동일한 영역을 다시 방문하는지 밝히는 것이며, 다음과 같이 공식화할 수 있다:

$$[0045] \quad RP_{ij} = \Theta(\varepsilon - \|\vec{s}_i - \vec{s}_j\|), \quad \vec{s}(\cdot) \in \mathbb{R}^t, \quad i, j = 1, \dots, t, \quad (1)$$

[0046] 여기서 RP_{ij} 는 RP의 픽셀이고, $\Theta(\cdot)$ 는 임계값 ε 를 통해 거리 행렬을 이진화하는 Heaviside 함수이며, $\vec{s}_i$ 는 위상 공간의 i 번째 상태이며, X 의 i 번째 위치에서 샘플링된 후속 단계에서 매핑된다. $\|\cdot\|$ 는 norm이고, T 는 총 상태 수이다. 그러나 RP는 경향 혼란의 문제를 가지고 있어 추가 적용을 제한한다.

[0047] MTF(Markov Transition Field)는 시간 도메인의 정보를 보존하면서 이산화된 시계열에 대한 전이 확률 필드를 나타낸다. MTF를 구축하기 위해 먼저 분위수를 기반으로 X 의 Q bins를 정의하고 모든 $x_i \in X$ 를 해당 bin q_j ($j \in [1, Q]$)에 할당한다. 그런 다음 시간 축을 따라 1차 마르코프 체인 방식으로 분위수 bin 사이의 전이를 계산하여 $Q \times Q$ 가중치 인접 행렬을 구성한다. 행렬은 $\sum_j v_{ij}$ 에 의해 정규화되며, 여기서 v_{ij} 는 $q_i \rightarrow q_j$ 의 전이 확률을 나타낸다. 그러나 이 행렬은 공간 영역 특성의 비상대성으로 인해 정보가 손실되는 문제가 있다. 정보 손실을 처리하기 위해, 행렬은 공간 도메인 위치에 해당하는 전이 가능성을 확장하여 MTF로 변환된다. MTF 행렬은 다음과 같이 정의된다:

$$[0048] \quad MTF = \begin{bmatrix} v_{i,j|x_1 \in q_i, x_1 \in q_j} & \dots & v_{i,j|x_1 \in q_i, x_t \in q_j} \\ v_{i,j|x_2 \in q_i, x_1 \in q_j} & \dots & v_{i,j|x_2 \in q_i, x_t \in q_j} \\ \vdots & \ddots & \vdots \\ v_{i,j|x_t \in q_i, x_1 \in q_j} & \dots & v_{i,j|x_t \in q_i, x_t \in q_j} \end{bmatrix} \quad (2)$$

[0049] 본 발명의 실시예에서는 데이터가 0과 1 사이에서 정규화되어 16개의 bins를 사용했다. 그럼에도 불구하고, MTF는 과적합의 위험을 포함하는 역 맵의 불확실성을 가지고 있다.

[0050] GAF(Gramian Angular Field)는 데카르트 좌표 대신 극좌표계에서 시계열을 나타낸다. 특히 X 는 모든 값이 구간 $[-1, 1]$ 내에 포함되도록 먼저 크기가 조정된다. 그런 다음 값을 각도 코사인으로 인코딩하여 재조정된 X 를 나

타내는 X' 을 극좌표로 나타내고, 식은 다음과 같다:

$$\begin{cases} \varphi = \arccos(x'_i) \\ r = \frac{\mathfrak{T}_i}{L} \end{cases} \quad (3)$$

여기서 x'_i 는 X' 의 i 번째 샘플이고, $\mathfrak{T}_i$ 는 타임스탬프이며, L 은 극좌표계의 스펠을 정규화하는 상수 계수이다. GAF는 R을 통해 시간적 의존성과 상관관계를 보존할 수 있다. X' 를 극좌표계로 변환한 후, GAF는 GASF와 GADF로 구분된다. 서로 다른 시간 간격 내에서 시간적 상관관계를 식별하는 각 지점은 합산되거나 차감되고, 식은 다음과 같다:

$$\begin{cases} GASF = \cos(\varphi_i + \varphi_k) \\ GADF = \cos(\varphi_i - \varphi_k) \end{cases} \quad (4)$$

도 3은 본 발명의 일 실시예에 따른 TCA를 설명하기 위한 도면이다.

도 3을 참조하면, F_{ϕ_l} 이 E_{ϕ} 의 1번째 레이어의 출력으로 주어지면 bi-GRU는 각 방향의 컨텍스트를 통합하기 위해 F_{ϕ_l} 의 각 행과 열에 대해 배치된다. h 의 채널은 픽셀 간의 시간적 관련성을 계산하기 위해 C' 로 축소된다. F_{ϕ_l} 의 각 픽셀은 $\alpha_{i,j}$ 에 의해 가중치 합산됩니다.

본 발명의 실시예에 따르면, 이전 모듈의 출력에 주의를 기울여 특징 추출부의 다음 모듈로 전달하는 TCA를 제안한다. ReNet에서는 반복 신경망(RNN)이 이미지 전반에 걸쳐 네 방향으로 배치되어 특징을 추출한다. ReNet에서 영감을 받은 TCA는 픽셀 간의 정보 관련성을 캡처하기 위해 특징 맵의 각 행과 열을 따라 bi-GRU를 배치한다. 이 주의 맵 생성은 각 픽셀에 대한 시간 단계 간의 시간적 상관 관계를 포함하는 인코딩된 이미지에 효과적이다. 게다가, attended 픽셀은 네트워크가 쓸모없는 특징을 피하면서 정보를 제공하는 컨텍스트만 학습할 수 있도록 한다 [46].

도 3에 나타난 바와 같이, $F_{\phi_l} \in \mathbb{R}^{C \times H \times W}$ 는 가중치 ϕ_l 과 함께 특징 추출부 E_{ϕ} 의 1번째 층의 출력을 나타낸다(310). 여기서 C , H , W 는 채널 수, 높이, 폭을 나타내며, ϕ 는 E 의 모든 훈련 가능한 매개변수를 나타낸다. bi-GRU는 먼저 F_{ϕ_l} 의 각 행에 배치되어 각 방향의 컨텍스트를 통합한다. 그런 다음 채널 축을 따라 두 개의 숨겨진 상태가 연결되어 $h^R \in \mathbb{R}^{2M \times H \times W}$ 를 얻는다. 여기서 M 은 은닉 상태의 차원을 나타낸다. 다음으로, h^R 의 각 열에 또 다른 big-GRU를 배치하고 두 개의 출력 은닉 상태를 이전과 동일한 방식으로 연결하여 $h \in \mathbb{R}^{2M \times H \times W}$ 를 얻는다. h 에는 각 방향의 컨텍스트와 혼합된 글로벌 컨텍스트가 포함된다.

h 의 채널 수는 그룹화된 컨볼루션(320)을 사용하여 C' 로 축소되어 픽셀 간의 관련성을 계산한다. 여기서 $C' = H \times W$ 이다. 이후, h 의 각 픽셀은 주의 가중치 $\alpha_{i,j} \in \mathbb{R}^{C'}$ 를 생성하기 위해 소프트맥스 함수를 통해 정규화된다(330):

$$\alpha_{i,j}^U = \frac{\exp(h_{i,j}^U)}{\sum_{V=1}^{C'} \exp(h_{i,j}^V)} \quad (5)$$

[0061] 여기서 h_{ij}^U 는 $h, i \in \{1, \dots, W\}$, $j \in \{1, \dots, H\}$, $U \in \{1, \dots, C'\}$ 의 U 번째 값이며, α_{ij}^U 는 α_{ij} 의 U 번째 값이고 참조된 픽셀과의 관련성을 포함한다. 마지막으로, F_{ϕ_l} 의 각 픽셀은 잔차 주의 메커니즘에 기초하여 $1 + \alpha_{ij}$ 로 가중된다:

$$[0062] F'_{\phi_l} = (1 + \alpha) \otimes F_{\phi_l} \quad (6)$$

[0063] 여기서, F_{ϕ_l} 은 주의 특징 맵, α 는 주의 맵, $\otimes$ 는 요소별 곱을 나타낸다.

[0065] 도 4는 본 발명의 일 실시예에 따른 MFF의 전체 기능 개선 프로세스를 설명하기 위한 도면이다.

[0066] 본 발명의 일 실시예에 따른 교차 모달리티 특징은 먼저 f_i 에 $\hat{f}$ 의 다른 세 가지 특징을 쌍으로 곱하여 얻을 수 있다(410). f_i^{cm} 은 교차 모달리티 특징을 평균화(420)하여 구하고, 훈련 가능한 매개변수 ω_γ^i 및 ω_β^i 를 사용하여 θ_i 를 생성하는 데 사용된다(430). 여기서 $\sigma(\cdot)$ 는 활성화 함수이다. 마지막으로 f_i 는 θ_i 에 의해 정제되고, 위의 과정은 $\hat{f}$ 의 각 특징에 적용된다.

[0067] 제안된 MFF는 특징 정제와 융합의 두 단계로 구성된다. MFF는 교차 모달리티 특징과 적응형 커널을 사용하여 각 특징을 차별적으로 정제한다.

[0068] 입력 인코딩된 이미지 $\hat{I} = \{I_1, I_2, I_3, I_4\}$ 에 대해 각 이미지의 특징은 E_ϕ 를 사용하여 얻는다:

$$[0069] \hat{f} = \{f_1, f_2, f_3, f_4\}. \quad (7)$$

[0070] 그런 다음 f_i 를 E_ϕ 을 사용하여 추출한 특징 중 하나로 정의한다. 여기서 $f_j \in \hat{f}$ 이다. 도 4에 도시된 바와 같이 f_i 에 대한 교차 모달리티 특징을 얻기 위해 먼저 f_i 를 $\hat{f}$ 쌍의 다른 세 가지 특징에 곱한다. 이러한 교차 모달리티 특징은 행별 평균으로 집계된다:

$$[0071] f_i^{cm} = \frac{1}{3} \sum_{\substack{j=1 \\ (i \neq j)}}^4 (f_i \otimes f_j) \quad (8)$$

[0072] 여기서 $f_j \in \hat{f}$ 이고, f_i^{cm} 는 집계된 교차 모달리티 특징이며, $\otimes$ 는 요소별 곱을 나타낸다. 다음으로, 커널 θ_i 는 f_i^{cm} 와 훈련 가능한 매개 변수, 즉 $\omega_\beta^i \in \mathbb{R}^{\frac{m}{\rho} \times m}$ 와 $\omega_\gamma^i \in \mathbb{R}^{m \times \frac{m}{\rho}}$ 를 사용하여 생성된다.

$$[0073] \theta_i = \omega_\gamma^i (\sigma(\omega_\beta^i (f_i^{cm}))) \quad (9)$$

[0074] 여기서 ρ 는 차원 감소 초매개변수이며, $\sigma(\cdot)$ 는 ReLU 함수를 의미한다. 이 커널은 모델이 특징 간의 상관관계를 이해할 수 있도록 하며, ReLU의 비선형성에 의한 유연한 변환으로 이어진다. 이제 f_i 는 f_i 와 θ_i 를 곱하여 정제된다:

$$[0075] f_i' = f_i \otimes \theta_i \quad (10)$$

[0076] 여기서 f'_i 는 f_i 의 정제된 특징이다. 상술된 단계를 $\hat{f}$ 의 각 특징에 적용한다. 마지막으로, 정제된 특징인 $\{f'_1, f'_2, f'_3, f'_4\}$ 이 평균화되고, 결과는 입력 시계열의 최종 특징으로 간주된다:

$$z = \frac{1}{4} \sum_{i=1}^4 f'_i \quad (11)$$

[0077] 여기서 z 는 입력 시계열의 최종 특징인 MFF의 출력이다. 본 발명의 실시예에 따른 커널을 생성하는 프로세스는 쿼리 데이터의 분류 오류를 최소화하여 훈련된다.

[0079] 쿼리 데이터를 분류하려면 NPC를 선택한다. 지표 기반 FSL 접근 방식이 단순하고 성능이 좋기 때문에 널리 사용되는 반면, 메타 학습 접근 방식은 기본 클래스와 보이지 않는 클래스 간에 큰 도메인 이동이 존재할 때 적응하는 데 어려움이 있다.

[0080] 임베딩 함수 $g_{\phi, \omega}(\cdot) : \mathbb{R}^t \rightarrow \mathbb{R}^m$ 을 정의하자. 이는 TCA와 MFF를 통해 특징 벡터를 계산한다. 여기서 ϕ 과 ω 은 각각 특징 추출부 E_{ϕ} 와 MFF에 대한 훈련 가능한 매개 변수이다. $g_{\phi, \omega}(\cdot)$ 을 사용하여 클래스 k 의 특징, $\{z_0^k, \dots, z_{N_k}^k\}$ 을 얻을 수 있다. 여기서 $z_i^k = g_{\phi, \omega}(X_i^k)$ 이고, N_k 는 클래스 k 의 레이블이 있는 지원 샘플의 수를 나타낸다. 다음으로 NPC는 N 클래스에 대한 클래스 프로토타입 $\hat{P} = \{p_1, \dots, p_{N-1}\}$ 을 계산한다:

$$p_k = \frac{1}{N_k} \sum_{i=1}^{N_k} z_i^k \quad (12)$$

[0082] 여기서 p_k 는 클래스 k 의 프로토타입이다. 그런 다음 분류부는 $z^* = g_{\phi, \omega}(X^*)$ 에 대한 클래스에 대해 소프트맥스와 같은 레이블 분포를 생성하며, 이는 임베딩 공간에서 쿼리 샘플 X^* 의 특징을 나타낸다.

$$\mathbb{P}_{(\phi, \omega)}(y = k | z^*) = \frac{\exp(-d(z^*, p_k))}{\sum_{c=1}^N \exp(-d(z^*, p_c))} \quad (13)$$

[0084] 여기서 $d(\cdot)$ 는 거리 지표이며, 이는 본 발명의 실시예에 따른 유클리드 거리이고, $d(z^*, p_k)$ 는 z^* 와 p_k 사이의 거리이다. 학습 프로세스는 실제 클래스 k 에 따라 음의 로그 확률 $\mathcal{L} = -\log_{\mathbb{P}_{\phi, \omega}}(y = k | z^*)$ 을 최소화함으로써 진행된다. 자세한 학습 절차는 알고리즘 2에 나와 있다.

[0086] 도 5는 본 발명의 일 실시예에 따른 시계열 데이터 분류를 위한 데이터 융합 방법을 설명하기 위한 흐름도이다.

[0087] 본 발명의 실시예에 따른 시계열 데이터 분류를 위한 데이터 융합 방법은 특징 추출부를 통해 입력 이미지에 대해 TCA(Temporal-Context Attention)를 이용하여 특징 맵의 글로벌 컨텍스트를 통합하고 다른 픽셀과 정보 의존성을 갖는 해당 픽셀을 강조하여 특징을 추출하는 단계(510), 특징 융합부를 통해 상기 추출된 특징에 대해 MFF(Meta Feature Fusion)를 이용하여 교차 모달리티 특징을 기반으로 커널을 생성하고, 생성된 커널을 이용하여 각 특징을 세분화하고 정제된 특징을 융합하는 단계(520) 및 분류부를 통해 상기 융합된 특징에 대하여 임베딩 공간에서 가장 가까운 클래스 프로토타입으로 분류하는 단계(530)를 포함한다.

[0088] 단계(510)에서, 특징 추출부를 통해 입력 이미지에 대해 TCA(Temporal-Context Attention)를 이용하여 특징 맵의 글로벌 컨텍스트를 통합하고 다른 픽셀과 정보 의존성을 갖는 해당 픽셀을 강조하여 특징을 추출한다.

[0089] 본 발명의 실시예에 따르면, 상기 입력 이미지를 인코딩하기 위해 TCA를 이용하여 입력 이미지의 픽셀 간 정보

의존성을 캡처하기 위한 bi-GRU(bi-directional gated recurrent unit)를 통해 글로벌 컨텍스트를 통합하고, FSL(Few-shot learning)을 통해 시계열의 인코딩된 이미지를 훈련시킨다.

- [0090] 단계(520)에서, 특징 융합부를 통해 상기 추출된 특징에 대해 MFF(Meta Feature Fusion)를 이용하여 교차 모달리티 특징을 기반으로 커널을 생성하고, 생성된 커널을 이용하여 각 특징을 세분화하고 정제된 특징을 융합한다.
- [0091] 본 발명의 실시예에 따르면, 상기 추출된 서로 다른 복수의 특징을 쌍으로 곱하여 교차 모달리티 특징을 획득하고, 상기 교차 모달리티 특징을 평균화한 후 상기 교차 모달리티 특징과 적응형 커널을 이용하여 각 특징을 차별적으로 정제하며, 상기 정제된 각 특징을 평균화하여 최종 특징을 출력한다.
- [0092] 단계(520)에서, 상기 분류부를 통해 상기 융합된 특징에 대하여 임베딩 공간에서 가장 가까운 클래스 프로토타입으로 분류한다.
- [0093] 본 발명의 실시예에 따르면, 상기 특징 추출부 및 상기 특징 융합부를 통해 특징 벡터를 계산하고, 미리 정해진 임베딩 함수를 이용하여 클래스 특징을 획득하여, NPC(Nearest Prototype Classifier)를 통해 클래스에 대한 클래스 프로토타입을 계산한다.
- [0094] 본 발명의 실시예에 따르면, 임베딩 공간에서 쿼리 샘플의 특징을 나타내는 각 클래스에 대한 레이블 분포를 생성하며, 클래스 특징 및 클래스 프로토타입 간의 유클리드거리에 따라 해당 클래스 특징을 가까운 클래스 프로토타입으로 분류한다.
- [0096] 도 6은 본 발명의 일 실시예에 따른 메타 특징 융합 알고리즘을 나타내는 도면이다.
- [0097] 본 발명의 실시예에 따른 커널을 생성하는 프로세스는 쿼리 데이터의 분류 오류를 최소화하여 훈련된다. 자세한 절차는 도 6의 알고리즘 1과 같다.
- [0099] 도 7은 본 발명의 일 실시예에 따른 FSL에서 네트워크를 훈련시키는 알고리즘을 나타내는 도면이다.
- [0100] 본 발명의 실시예에 따른 학습 프로세스는 실제 클래스 k 에 따라 음의 로그 확률 $\mathcal{L} = -\log_{\mathbb{P}_{\phi, \omega}}(y = k|z^*)$ 을 최소화함으로써 진행된다. 자세한 학습 절차는 도 7의 알고리즘 2와 같다.
- [0102] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.
- [0103] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embodiment)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이

터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

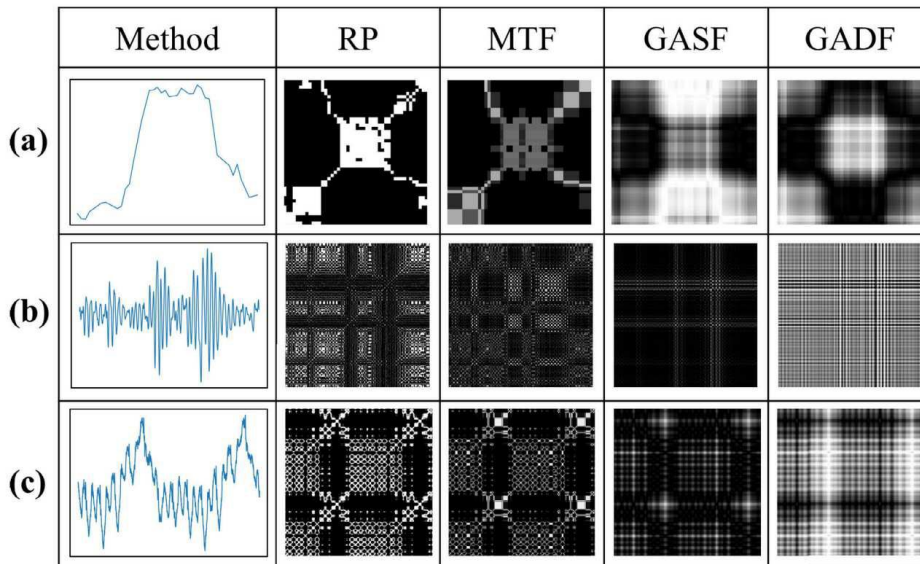
[0104] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

[0105] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

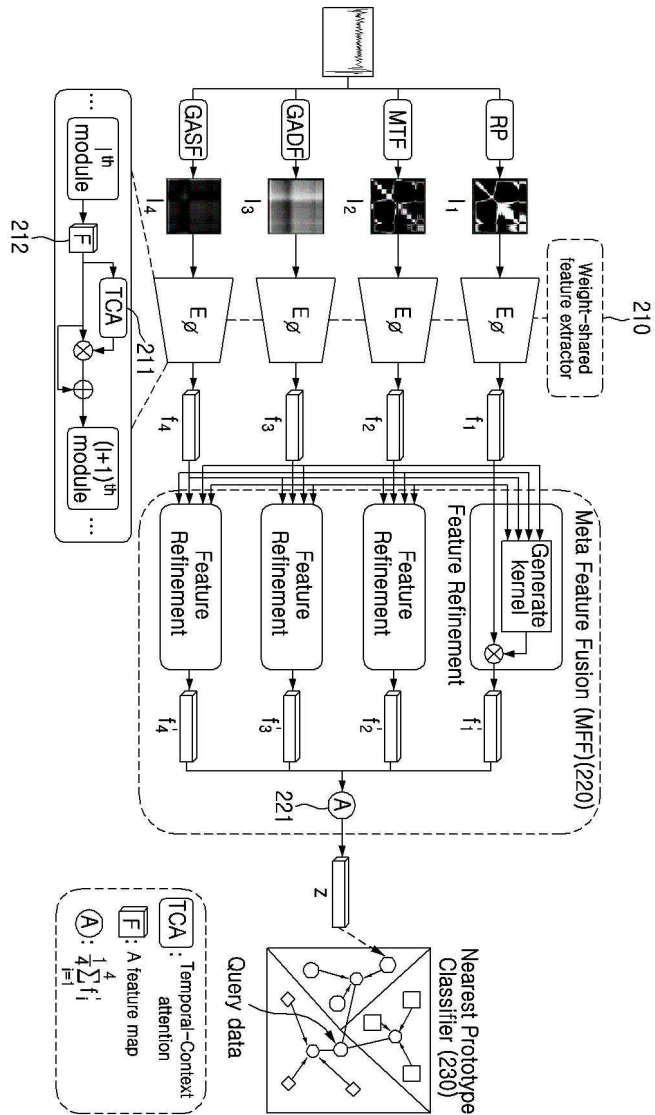
[0106] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

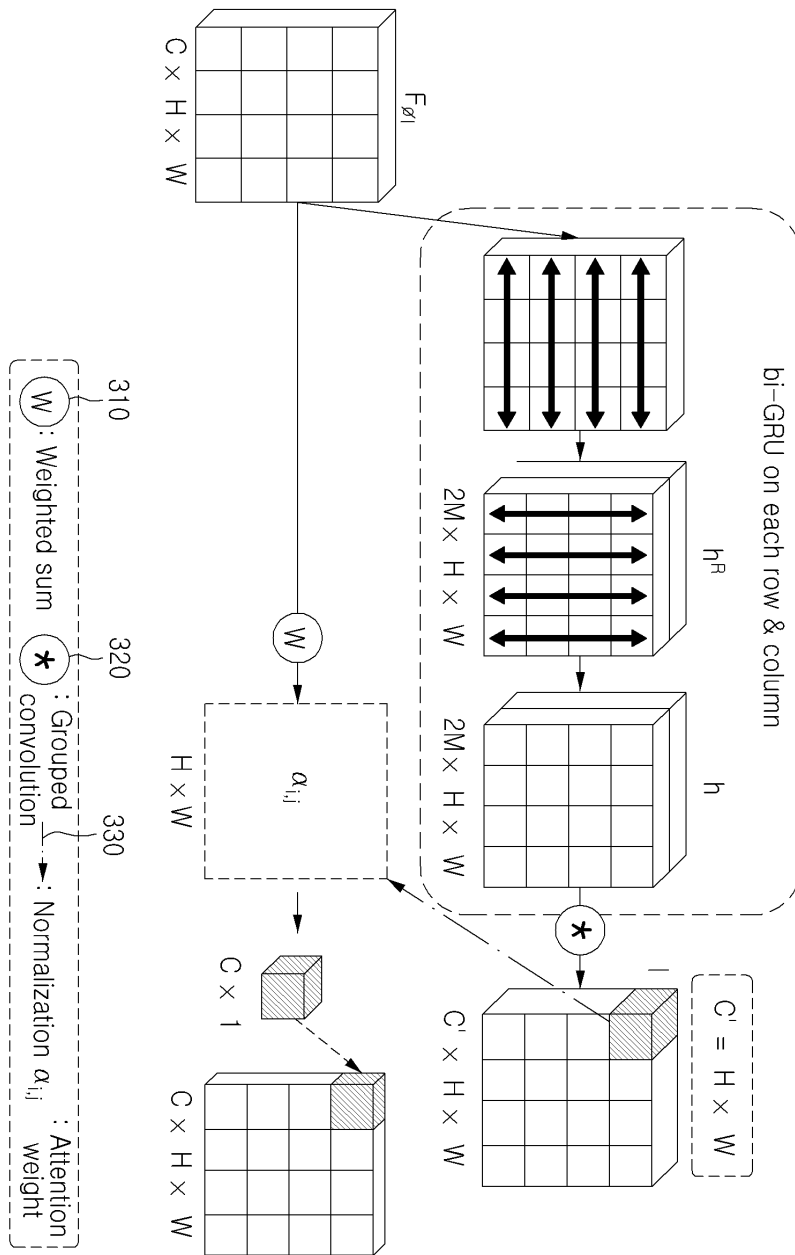
도면1



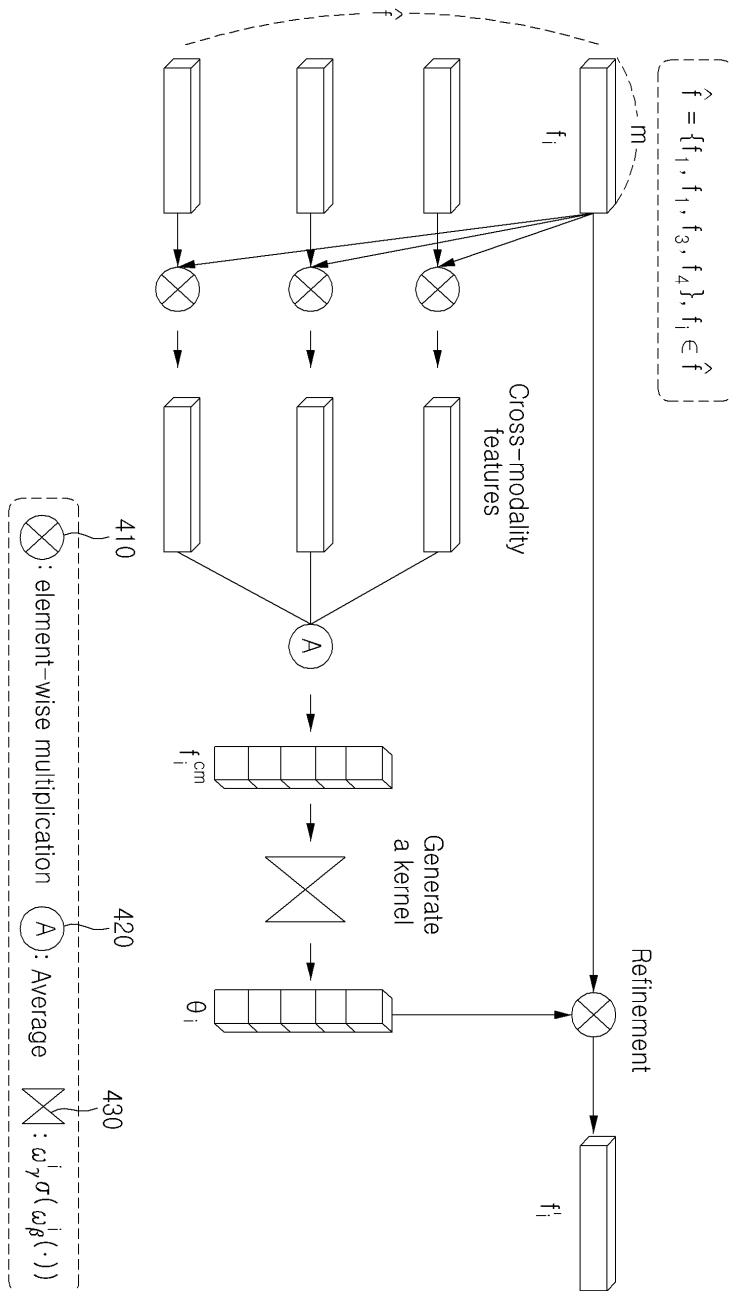
도면2



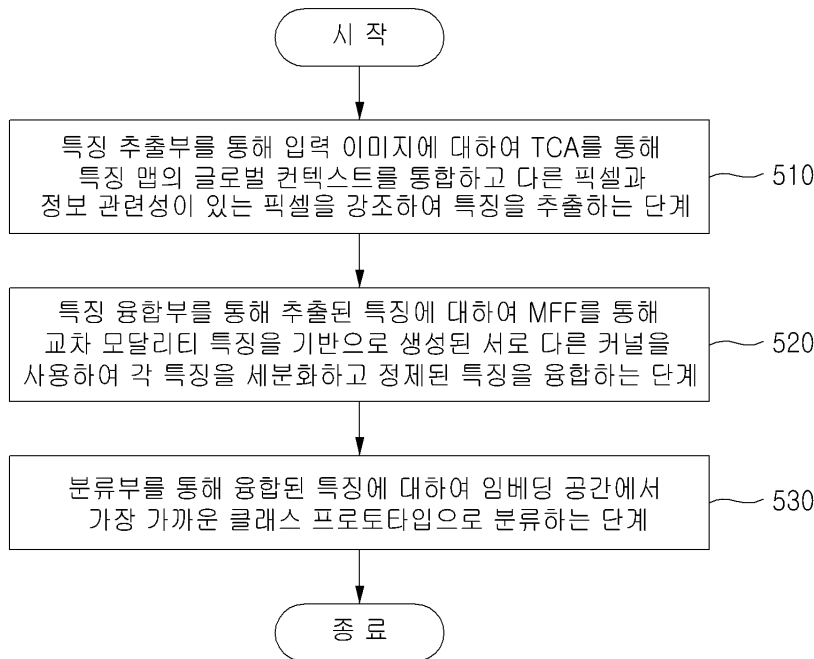
도면3



도면4



도면5



도면6

Algorithm 1 Procedure of MFF. $\otimes$ Is Element-Wise Multiplication. ω_{β}^i and ω_{γ}^i Are Trainable Parameter for Generating the Kernel θ_i . $\sigma(\cdot)$ Refers to the ReLU Function

Input: $\hat{f} = \{f_1, f_2, f_3, f_4\}$.

Output: A feature z .

1: for $i = 1, 2, 3, 4$ do

2: $f_i^{cm} \leftarrow \frac{1}{3} \sum_{\substack{j=1 \\ (i \neq j)}}^4 (f_i \otimes f_j)$

3: $\theta_i \leftarrow \omega_{\gamma}^i (\sigma(\omega_{\beta}^i (f_i^{cm})))$ ▷ Generate a kernel

4: $f_i' \leftarrow f_i \otimes \theta_i$ ▷ Refine a feature

5: end for

6: $z \leftarrow \frac{1}{4} \sum_{i=1}^4 f_i'$ ▷ Fuse the refined features

도면7

Algorithm 2 Process of Training the Network on FSL Procedure. N , N_k , N_Q Are the Number of Classes per Task, Support Examples Labeled With Class k , and Unlabeled Query Examples, Respectively. ϕ and ω Denotes All Trainable Parameters for the Feature Extractor E_ϕ and MFF Respectively. $g_{\phi, \omega}(\cdot)$ Is an Embedding Function to Compute the Feature Vector via TCA and MFF.

Input: Training set D_{train} .

Output: The loss L for a sampled task.

```

1: Select class indices containing an task
2: for  $k$  in  $\{1, \dots, N\}$  do
3:   Select support samples,  $S_k^\tau = \{X_i, Y_i\}_{i=1}^{N_k}$ 
4:   Select query samples,  $Q_k^\tau (X_i^*, Y_i^*)_{i=1}^{N_Q}$ 
5:   for  $i$  in  $\{1, \dots, N_k\}$  do
6:      $z_i \leftarrow g_{\phi, \omega}(X_i)$ 
7:   end for
8:    $p_k = \frac{1}{3} \sum_{i=1}^{N_k} z_i$ 
9: end for
10:  $L \leftarrow 0$ 
11: for  $k$  in  $\{1, \dots, N\}$  do
12:   for  $(X^*, Y^*)$  in  $Q^\tau$  do
13:      $z^* \leftarrow g_{\phi, \omega}(X^*)$ 
14:     Compute  $-\log_{P_{\phi, \omega}}(y = k|z^*)$ 
15:     Update  $L$ 
16:   end for
17: end for

```
